

**INTERLINKED POPULATION BALANCE AND
CYBERNETIC MODELS
FOR THE SIMULTANEOUS
SACCHARIFICATION AND FERMENTATION OF
NATURAL POLYMERS**

HO YONG KUEN

**FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2015

**INTERLINKED POPULATION BALANCE AND
CYBERNETIC MODELS
FOR THE SIMULTANEOUS
SACCHARIFICATION AND FERMENTATION OF
NATURAL POLYMERS**

HO YONG KUEN

**THESIS SUBMITTED IN FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF
PHILOSOPHY**

**FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2015

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: *Ho Yong Kuen* (I.C./ Passport No: *840309-13-5117*)

Registration/ Matric No.: *KHA 110112*

Name of Degree: *Doctor of Philosophy*

Title of Project/ Paper/ Research Report/ Dissertation/ Thesis ("this Work"):
*Interlinked Population Balance and Cybernetic Models for The Simultaneous
Saccharification and Fermentation of Natural Polymers*

Field of Study: *Bioprocess Engineering*

I do solemnly and sincerely declare that:

- (1) I am the sole author/ writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether I intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's signature

Date:

Subscribed and solemnly declared before,

Witness's signature

Date:

Name:

Designation:

ABSTRACT

The generation of important and useful products (e.g. ethanol, lactic acid etc.) through microbial fermentation often involves the breakdown of complex polymeric feedstock such as starch and cellulose through enzymatic scissions followed by subsequent metabolic conversion. The interplay between the kinetics of enzymatic depolymerization and the response of the microbes towards changes in the abiotic phase is critical for the adequate description of such a complex process. In this work, two unrelated frameworks, i.e. the Population Balance Modelling (PBM) and the Cybernetic Modelling (CM) were interlinked to model such a system. Specifically, the PBM technique was used to describe the enzymatic depolymerization whereas the CM framework was used to model the microbial response toward complex environmental changes. As the enzymes required to break down polymeric substrates are produced by the microbes, a more general treatment of the secretion of extracellular enzyme was also proposed in the CM model. In the course of interlinking the two frameworks, the numerical techniques for solving Population Balance Equations (PBEs) were explored. In this regard, the Fixed Pivot (FP) technique was successfully modified to solve chain-end scission which resembles the action of enzyme which removes a monomer from the end of a polymer chain. This method was further extended to include random scission (resembling the action of enzyme which randomly hydrolyzes the bond of a polymer chain) and mixed scission involving both modes. Simulation results showed that the FP technique was able to solve chain-end scission and simultaneous random and chain-end scissions to a high degree of accuracy using 0.02% and 1.2% of the time required for solving the exact case respectively. One notable feature of the interlinked framework is the flexible linkage, which allows the individual PBM and CM components to be independently modified to the desired levels of detail. The interlinked PBM and CM framework was implemented on two case studies

involving the Simultaneous Saccharification and Fermentation (SSF) of starch by two recombinant yeast strains capable of excreting glucoamylase alone or together with α -amylase. The simulation results revealed that the proposed framework captured features not attainable by existing approaches. Examples of such include the ability of the model to indicate (in case study one) that an appropriate amount of glucose (7 g) mixed with starch (30 g) as initial substrates yielded an optimum productivity of ethanol. Not only that, the model showed (in case study two) that SSF is indifferent to the type of starch when both enzymes are present as opposed to when only glucoamylase is present, where the time required for ethanol concentration to peak differed by more than 30 hours between different starches. Thus, the effect of various enzymatic actions on the temporal evolution of the polymer distribution and how the microbes respond to the initial molecular distribution of the polymers can be studied. Such a framework also enables a more molecular and fundamental study of a complex SSF system, a feat which heretofore was unattainable by existing SSF models.

ABSTRAK

Pengeluaran hasil penting dan berguna (misalnya etanol, asid laktik dan lain-lain) melalui penapaian mikrob sering melibatkan pecahan bahan mentah polimer kompleks seperti kanji dan selulosa melalui pengguntingan enzim diikuti dengan penukaran metabolik berikutnya. Hubungan di antara kinetik enzim penyahpolimeran dan tindak balas mikrob terhadap perubahan dalam fasa abiotik adalah kritikal untuk menerangkan suatu proses yang kompleks sebegini. Dalam karya ini, dua pendekatan yang tidak berkaitan, iaitu Permodelan Imbangan Populasi (PIP) dan Permodelan Ciberetik (PC) telah dikaitkan untuk memodel sistem seperti ini. Secara khusus, teknik PIP digunakan untuk menerangkan penyahpolimeran enzim manakala rangka kerja PC digunakan untuk memodel tindak balas mikrob terhadap perubahan persekitaran yang kompleks. Oleh sebab enzim yang diperlukan untuk penyahpolimeran dirembeskan oleh mikrob sendiri, satu olahan yang lebih umum untuk rembesan enzim ke luar sel juga dicadangkan dalam model PC. Dalam proses mengaitkan kedua-dua rangka kerja ini, kaedah berangka untuk menyelesaikan Persamaan PIP untuk pengguntingan enzim telah diterokai. Dalam hal ini, teknik Pangsi Tetap (PT) telah diubahsuai dengan sewajarnya untuk menyelesaikan pengguntingan akhir rantai yang menyerupai tindakan enzim yang membuang satu monomer dari hujung rantai polimer. Rangka kerja penyelesaian ini telah juga digunakan untuk menyelesaikan pengguntingan rawak (menyerupai tindakan enzim yang secara rawak memotong ikatan rantai polimer) dan campuran kedua-dua pengguntingan akhir rantai dan rawak. Keputusan simulasi menunjukkan bahawa teknik PT itu dapat menyelesaikan pengguntingan akhir rantai dan campuran kedua-dua jenis pengguntingan akhir rantai dan rawak dengan ketepatan yang tinggi dan hanya menggunakan 0.02% dan 1.2% daripada masa yang diperlukan untuk menyelesaikan sistem persamaan penuh masing-masing. Satu ciri yang penting dalam model yang terhubungkait ini adalah

perkaitan yang mudah alih, seterusnya membolehkan komponen individu PIP dan PC untuk bebas diubahsuai kepada tahap keperincian yang dikendaki. Rangka kerja PIP-PC ini telah dilaksanakan ke atas dua kajian kes yang melibatkan Pemanisan dan Penapaian Serentak (PPS) kanji oleh dua jenis yis rekombinan yang mampu merembes glukoamilas atau kedua-dua glukoamilas dan α -amilas. Keputusan simulasi menunjukkan bahawa rangka kerja yang dicadangkan mempunyai ciri-ciri yang tidak dapat dicapai dengan pendekatan yang sedia ada. Antara contoh-contoh itu termasuk keupayaan model (dalam kajian kes pertama) untuk menunjukkan bahawa 7 g glukosa dicampur dengan 30 g kanji sebagai bahan mentah awal akan mencapai produktiviti etanol yang terbaik. Bukan itu sahaja, model menunjukkan (dalam kajian kes kedua) bahawa PPS tidak bergantung kepada jenis kanji apabila kedua-dua enzim hadir. Apabila hanya glukoamilas hadir, masa yang diperlukan untuk kepekatan etanol mencapai ke tahap maksimum berbeza sebanyak lebih daripada 30 jam antara kanji yang berbeza. Justeru itu, kesan pelbagai tindakan enzim pada perubahan dengan masa taburan polimer boleh dikaji. Rangka kerja sedemikian juga membolehkan kajian yang lebih molekular dan bersifat asas untuk sistem PPS yang kompleks, satu pencapaian yang melangkaui semua model PPS yang sedia ada.

ACKNOWLEDGMENTS

First and foremost, I thank my God and my Lord Jesus Christ, who abundantly supplies fresh grace into my life daily, without whom I would never even have dared to imagine the completion of this thesis. May my life always be a testimony of His grace and love for humanity. To Him be the glory in whatever that I do. Amen.

During this entire period of my candidacy, a few people had greatly contributed to the birth of this thesis in one way or another. The first being my academic advisor and friend, Dr. Yeoh Hak Koon from the Department of Chemical Engineering (University of Malaya, Malaysia). During our regular sessions, because of his highly meticulous character, a great amount of time was spent in going through the minute details of this work. His constant attempts at achieving what I deem as perfection often resulted in ‘excessive’ revisions of a single figure. As I reflect on myself today, my attitude towards research is definitely the result of his cultivation. For all that, I sincerely thank him for the willingness to dot the i’s and to cross the t’s with me on this project. If someone were to ask me, I would not hesitate to recommend him as the most erudite academic advisor in the department.

I would also like take this opportunity to thank my second academic advisor, Dr. Ngoh Gek Cheng from the Department of Chemical Engineering (University of Malaya, Malaysia). Dr. Ngoh is to me a very motherly figure in the department. A short chat with her would always leave me feeling hopeful about my research. Because of her constant support, my engagements with the departmental facilities proceeded smoothly without any obstructions. Moreover, her wide knowledge of how the organization works had ensured that this project always secured the required resources.

Many thanks also to Dr. Pankaj Doshi from the National Chemical Laboratory, India. The idea of interlinking the population balance and the cybernetic modelling framework came primarily during our discussion with Dr. Pankaj. Despite the differences in the time zone, Dr. Pankaj has never failed to show up in our teleconferencing sessions. Whenever I ran out of ideas, Dr. Pankaj was always able to offer a more elegant and fresh way of viewing the problems at hand. His contributions in laying the fundamental grounds of this thesis can never be ignored.

Given this precious opportunity, I would also like to express my gratitude to my dad, my mum and my brother, who have all along without fail granted me their unconditional love and support. Their persistence in showering me with their care and affections enabled me to look beyond the daily obstacles and challenges and strive towards the accomplishment of this work. Special thanks also to my pastor Peter Sze, who always encourages me during the difficult moments of life, offering me fresh bread from the throne of grace. To my friends, thank you for rocking through life with me!

Last but not least, I would like to thank the University of Malaya for the financial support received through the Bright Sparks program, the PP Grant (PG091-2012B) and the UMRG Grant (RG134-11SUS).

TABLE OF CONTENTS

ABSTRACT.....	iii
ABSTRAK.....	v
ACKNOWLEDGMENTS	vii
TABLE OF CONTENTS.....	ix
LIST OF FIGURES	xiii
LIST OF TABLES	xxi
LIST OF AND SYMBOLS AND ABBREVIATIONS.....	xxiii
LIST OF APPENDICES	xxxix
CHAPTER 1 : INTRODUCTION.....	1
1.1 Research Background	1
1.2 Research Objectives.....	8
1.3 Structure of the Thesis	8
CHAPTER 2 : LITERATURE REVIEW.....	10
2.1 Simultaneous Saccharification and Fermentation (SSF)	10
2.1.1 Modelling of SSF Processes.....	11
2.2 Population Balance Modelling (PBM).....	16
2.2.1 Random Scission.....	18
2.2.2 Chain-End Scission	22
2.2.3 Population Balance Modelling of Saccharification Processes	26
2.3 Cybernetic Modelling (CM).....	28
2.3.1 Cybernetic Modelling Equations	28

2.3.2	Cybernetic Modelling for SSF Processes.....	32
CHAPTER 3 : RESEARCH METHODOLOGY		34
3.1	Preamble	34
3.2	General Workflow	35
3.3	Development of the Numerical Technique for Solving Population Balance Equations	36
3.3.1	Identification of the Relevant Stoichiometric Kernels, Rate Kernels and Initial Distribution for Solving the PBEs	36
3.3.2	Code Validation	38
3.3.3	Validation with the Exact (or Fully Discrete) Solution	39
3.4	Development of the Interlinked PBM and CM Framework	40
3.4.1	Identification of Relevant Stoichiometric Kernels, Rate Kernels and Initial Distribution for Solving PBEs	41
3.4.2	Identification of Model Parameters.....	41
CHAPTER 4 : MODELLING POLYMERIC SCISSIONS USING THE FIXED PIVOT TECHNIQUE		46
4.1	Modelling Chain-End Scission	46
4.1.1	Fully Discrete (Exact) Solution for Chain-End Scission	46
4.1.2	Fixed Pivot Discretization for Chain-End Scission	47
4.1.3	Meshing and Implementation.....	49
4.1.4	Case Study on Chain-End Scission.....	53
4.1.5	Observations of the FP Method for Solving Chain-End Scission.....	67
4.1.6	Guidelines for Meshing.....	73

4.2	Modelling Simultaneous Random and Chain-End Scissions.....	75
4.2.1	Fully Discrete (Exact) Solution for Simultaneous Random and Chain-End Scissions.....	76
4.2.2	Computing the Exact Solution for the Purpose of Validation.....	77
4.2.3	Fixed Pivot Discretization for Simultaneous Random and Chain-End Scissions.....	82
4.2.4	Case Study on Simultaneously Occurring Random and Chain-End Scissions	84
4.3	Observations of the FP Method for Solving Random Scission	96
4.3	Concluding Remarks.....	103
CHAPTER 5 : INTERLINKED POPULATION BALANCE AND CYBERNETIC MODELLING FRAMEWORK		104
5.1	Theoretical Framework.....	104
5.1.1	Population Balance Modelling for Enzymatic Scission.....	105
5.1.2	Cybernetic Modelling.....	108
5.1.3	Interlinked Population Balance and Cybernetic Framework	110
5.2	Case Study I: Growth of A Glucoamylase Producing Recombinant <i>S. cerevisiae</i> on Starch.....	114
5.2.1.	Model Formulation	114
5.2.2	Parameter Identification and Initial Conditions	121
5.2.3	Simulation Results	126
5.3	Case Study II: Growth of A Glucoamylase and α -amylase Producing Recombinant <i>S. cerevisiae</i> on Starch	140

5.3.1	Model Formulation	140
5.3.2	Parameter Identification and Initial Conditions	148
5.3.3	Simulation Results	153
5.4	Concluding Remarks.....	170
CHAPTER 6 : CONCLUSIONS, THESIS AND RECOMMENDATION.....		172
6.1	Conclusions and Thesis.....	172
6.2	Recommendations and Future Work.....	174
REFERENCES		176
LIST OF PUBLICATIONS AND CONFERENCES ATTENDED.....		195
APPENDICES		196

LIST OF FIGURES

Figure 1.1: Illustrating the linkage between the PBM and CM in a single framework. In particular, the induction/repression of the synthesis of extracellular depolymerase by the composition of the broth appears to be relatively unexplored within the CM framework.	6
Figure 2.1: Illustration of random scission where the cases 1 to 5 are the equally probable products resulting from the scission of a polymer chain with 6 monomer units and 5 bonds.	18
Figure 2.2: Illustration of chain-end monomer scission where monomer is removed successively from the end of the chain.	22
Figure 2.3: Simplified diagram illustrating the different components of the CM framework. The biotic phase is represented by the region enclosed within the dashed line.	31
Figure 3.1: Flowchart of the research workflow.....	35
Figure 3.2: The general workflow of nonlinear parameter estimation using the Genetic Algorithm (GA).	45
Figure 4.1: The discrete-continuous strategy with x_i being the pivot encompassed by v_i and v_{i+1} . The integer p is the number of pivots in the discrete region while q is the number of pivots in the continuous region, with $x_{p+q} = v_{p+q+1} = N$ ($N = \text{maximum DP}$).....	49
Figure 4.2: Transient of the mass concentration of monomer (glucose) using the exact and fixed pivot (FP) solutions for chain-end scission. Here $[p, q] = [100, 500]$, $r = 1.0109$ and $m_S(0) = 10$ g/L. The dimensionless time is normalized against the time required for 99% monomer production. The error in the initial mass due to discretization $[\varepsilon_S, \text{cf. Eq. (3.9)}]$ is $2.61 \times 10^{-4} m_S(0)$. The case of $[p, q] = [10, 30]$ with $r = 1.3007$ was given as a reference to coarsely resolved mesh.	56

Figure 4.3: Transients of the normalized mass concentration of oligomers using the exact and the fixed pivot (FP) solutions for chain-end scission. The highest peak corresponds to the DP7 oligomer, followed by the DP6 – DP2 oligomers in the order of decreasing peak heights. Here $[p, q] = [100, 500]$, $r = 1.0109$, $m_s(0) = 10$ g/L, and the error in the initial mass due to discretization (ε_s) is $2.61 \times 10^{-4} m_s(0)$. The dimensionless time is normalized against the time required for 99% monomer production.58

Figure 4.4: Transients of the molar concentration density using the exact and the fixed pivot (FP) solutions for chain-end scission. The dimensionless time (θ) is normalized against the time required for 99% monomer production. Here $[p, q] = [100, 500]$ and $r = 1.0109$. The exact solutions fall on the FP solutions for $DP = 1$59

Figure 4.5: Transient of the number-average DP using the exact and the fixed pivot (FP) solutions for chain-end scission. Here, $r = 1.0109$ for $[p, q] = [100, 500]$. The case of $[p, q] = [10, 30]$ with $r = 1.3007$ was given as a reference to coarsely resolved mesh. The dimensionless time is normalized against the time required for 99% monomer production.61

Figure 4.6: Transient of the weight-average DP using the exact and the fixed pivot (FP) solutions for chain-end scission. Here, $r = 1.0109$ for $[p, q] = [100, 500]$. The case of $[p, q] = [10, 30]$ with $r = 1.3007$ was given as a reference to coarsely resolved mesh. The dimensionless time is normalized against the time required for 99% monomer production.62

Figure 4.7: Transient of the polydispersity index using the exact and the fixed pivot (FP) solutions for chain-end scission. Here $[p, q] = [100, 500]$ and $r = 1.0109$. The case of $[p, q] = [10, 30]$ with $r = 1.3007$ was given as a reference to coarsely resolved mesh. The dimensionless time is normalized against the time required for 99% monomer production.63

Figure 4.8: Fraction of time required to produce the fixed pivot solution vs. the fraction of pivots used. Here, t_{FP} is the time required to obtain the fixed pivot solution, t_{Exact} is the time required to obtain the exact solution. 65

Figure 4.9: Effect of increasing the size of the exact problem according to $N = k \times 22496$ ($k = 1.0, 1.5, 2.0, 2.5$) while retaining the initial polydispersity index at 1.32 as well as using $[p, q] = [100, 500]$ for the fixed pivot (FP) solution. The geometric ratios employed were $r = 1.0109, 1.0117, 1.0123$, and 1.0128 in the order of increasing k . The molar concentration density is shown for the dimensionless time, $\theta = 1.08$. Here, $m_S(0) = 10$ g/L and the errors in the initial mass due to discretization (ϵ_S) were $2.6 \times 10^{-4} m_S(0)$, $1.7 \times 10^{-4} m_S(0)$, $1.2 \times 10^{-4} m_S(0)$, and $0.95 \times 10^{-4} m_S(0)$ in the increasing order of k 66

Figure 4.10: Performance of the fixed pivot (FP) solution in solving a relatively large chain-end scission problem, i.e. $N = 224,960$, by using $< 1\%$ of the number of equations employed by the exact solution. Here, the initial polydispersity index was 1.32, $[p, q] = [256, 1744]$, $r = 1.0039$, $m_S(0) = 100$ g/L and the error in the initial mass due to discretization (ϵ_S) was $0.26 \times 10^{-4} m_S(0)$. The molar concentration density is shown for the dimensionless time, $\theta = 1.08$ when slightly more than 99% of the monomers had been formed. Here t_{FP} and t_{Exact} are respectively the time required to produce the FP and the exact solutions. 67

Figure 4.11: Temporal evolution of the molar concentration density using the fixed pivot (FP) and the exact solutions for chain-end scission. The highest peak corresponds to $t = 0$ time units, followed by $t = 144, 300, 456$, and 600 time units in the order of decreasing peak heights. Here, $[p, q] = [10, 545]$ where p and q are the number of pivots for the discrete and continuous region respectively. 70

Figure 4.12: Illustrating the limitation of the fixed pivot technique in modelling chain-end scission for steep distributions. Here $v_m = 1$, x_i is the pivot for the i -th interval, and c_i

is the molar concentration density. The symbol c_{i+1}^* represents the linearly extrapolated concentration while \hat{c}_{i+1} represents the actual concentration at x_{i+1}^* 73

Figure 4.13: Minimum value of p (p_{min}) at a selected value of $p+q$ for a given N (solid line), with $v_m = 1$. The dashed lines correspond to $(p+q)/N$ 75

Figure 4.14: Illustrating the iterative procedure for obtaining $\mathbf{C}^{(i)}(t_k)$, $i = 1$ to w 81

Figure 4.15: Stiffness ratio ($= \max(|\lambda_i|) / \min(|\lambda_i|)$) evaluated from $i = 1$ to w , where λ_i is the vector of eigenvalues for $\mathbf{G}^{(i,i)}$, w is the total number of partitions for the molar concentration vector, and S is the size of each partitioned molar concentration vector. 87

Figure 4.16: Transient of the mass concentration of monomer (glucose) using the exact and fixed pivot (FP) solutions for simultaneous random and chain-end scissions. Here $[p, q] = [100, 500]$, $r = 1.0109$ and $m_S(0) = 10$ g/L. The dimensionless time is normalized against the time required for 99% monomer production and the error in the initial mass due to discretization [ε_s , cf. Eq. (3.9)] is $2.61 \times 10^{-4} m_S(0)$ 88

Figure 4.17: Transients of the normalized mass concentration of oligomers using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. The highest peak corresponds to the DP2 oligomer, followed by the DP3 – DP7 oligomers in the order of decreasing peak heights. Here $[p, q] = [100, 500]$, $r = 1.0109$, $m_S(0) = 10$ g/L, and the error in the initial mass due to discretization (ε_s) is $2.61 \times 10^{-4} m_S(0)$. The dimensionless time is normalized against the time required for 99% monomer production. 89

Figure 4.18: Transients of the molar concentration density using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. The dimensionless time (θ) is normalized against the time required for 99% monomer production. Here $[p, q] = [100, 500]$ and $r = 1.0109$. The exact solutions fall on the FP solutions for DP = 1. 91

Figure 4.19: Transient of the number-average DP using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. Here, $r = 1.0109$ for $[p, q] = [100, 500]$. The dimensionless time is normalized against the time required for 99% monomer production.	92
Figure 4.20: Transient of the weight-average DP using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. Here, $r = 1.0109$ for $[p, q] = [100, 500]$. The dimensionless time is normalized against the time required for 99% monomer production.	93
Figure 4.21: Transient of the polydispersity index using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. Here $[p, q] = [100, 500]$ and $r = 1.0109$. The dimensionless time is normalized against the time required for 99% monomer production.	94
Figure 4.22: Performance of the fixed pivot (FP) solution in solving a relatively large simultaneous random and chain-end scissions problem, i.e. $N = 224,960$, by using $< 1\%$ of the number of equations employed by the exact solution. Here, the initial polydispersity index was 1.32, $[p, q] = [256, 1744]$, $r = 1.0039$, $m_s(0) = 100$ g/L and the error in the initial mass due to discretization (ε_s) was $0.26 \times 10^{-4} m_s(0)$. The molar concentration density is shown for the dimensionless time, $\theta = 1.23$ when more than 99% of the monomers had been formed. Here t_{FP} and t_{Exact} are respectively the time required to produce the FP and the exact solutions. The exact solutions fall on the FP solutions for DP = 1.	95
Figure 4.23: Temporal evolution of molar concentration of the DP2 – DP7 oligomers using the fixed pivot (FP) and the exact solutions for random scission. The dimensionless time $\theta = 1$ corresponds to the time when more than 99% monomer had been formed. Here, $[p, q] = [10, 545]$ and $[p, q] = [100, 455]$ were used where p and q are the number of pivots for the discrete and continuous region respectively.	101

Figure 5.1: Simplified diagram illustrating the different components of the interlinked PBM and CM framework where the shaded region (excluding the excretion of extracellular depolymerase) is the standard CM framework. In this illustration, for simplicity the cell is assumed to excrete only one form of depolymerase for the interlinked framework.....	113
Figure 5.2: Transients of various quantities in the fermentation broth where model predictions are represented by lines and experimental data are represented by symbols. Here, starch (30 g/L) is the sole initial substrate.....	128
Figure 5.3: Transients of polymer properties corresponding to the system described in Figure 5.2. For the molar concentration density, the triangular symbol represents the concentration density of glucose where the lowest point corresponds to time = 10 h, followed by time = 20 h to time = 50 h in the order of increasing density values.....	130
Figure 5.4: Transients of various quantities in the fermentation broth when 3 g of glucose (dashed line), or 3 g of maltose (dashed-double-dotted line) was added in addition to 30 g of starch at time = 0 h. The profile for 30 g starch (solid line) as the sole substrate is given as reference.....	132
Figure 5.5: Transients of the level of key metabolic enzymes, plus the cybernetic variables U and V corresponding to the maltose-starch mixture presented in Figure 5.4.	134
Figure 5.6: Transients of various quantities in the fermentation broth when 30 g of maltotriose was added to 30 g of starch at time = 0 h.....	136
Figure 5.7: Effect of various mixture of initial substrates on the molar concentration density at 80 h.	137
Figure 5.8: Illustrating (a) the ethanol production with various mixtures of initial substrate and (b) the productivity of ethanol as a function of glucose addition, where A is the concentration of ethanol (subscripts 'p' and '0' denoting the peak and initial values	

respectively), m_1^{eq} is the total mass of substrate in the form of glucose equivalent (1 g starch = 1.11 g glucose), and t_p is the time required to achieve peak ethanol concentration.

..... 139

Figure 5.9: Transients of various quantities in the fermentation broth where model predictions are represented by lines and experimental data (Ülgen et al., 2002) are represented by symbols. Here, starch (38.4 g/L), glucose (0.79 g/L) and maltose (1.69 g/L) are initially present. Soluble starch with $[\overline{M}_n, \overline{M}_w, N, PD] = [160, 212, 878, 1.325]$ was used here. 155

Figure 5.10: Transients of the level of key metabolic enzymes, plus the cybernetic variables U and V corresponding to the simulation presented in Figure 5.9. 157

Figure 5.11: Transients of various quantities in the fermentation broth when only α -amylase is produced by the yeast. Soluble starch with $[\overline{M}_n, \overline{M}_w, N, PD] = [160, 212, 878, 1.325]$ was used here. 159

Figure 5.12: Transients of the level of key metabolic enzymes, plus the cybernetic variables U and V corresponding to the simulation presented in Figure 5.11 160

Figure 5.13: Transients of various quantities in the fermentation broth when only glucoamylase is produced by the yeast. Soluble starch with $[\overline{M}_n, \overline{M}_w, N, PD] = [160, 212, 878, 1.325]$ was used here. 162

Figure 5.14: Transients of the level of key metabolic enzymes, plus the cybernetic variables U and V corresponding to the simulation presented in Figure 5.13 163

Figure 5.15: Transients of various quantities in the fermentation broth when only glucoamylase is produced by the yeast. Here, dashed lines represent the results for starch A with $[\overline{M}_n, \overline{M}_w, N, PD] = [10000, 10100, 22496, 1.01]$, dashed double dotted lines represent the results for starch B with $[\overline{M}_n, \overline{M}_w, N, PD] = [4100, 5430, 22496, 1.324]$,

while solid lines represent the results for starch C with $\left[\overline{M}_n, \overline{M}_w, N, PD\right] = [160, 212, 878, 1.325]$ 166

Figure 5.16: Transients of various quantities in the fermentation broth when the yeast is capable of excreting both α -amylase and glucoamylase. Here, dashed lines represent the results for starch A with $\left[\overline{M}_n, \overline{M}_w, N, PD\right] = [10000, 10100, 22496, 1.01]$, dashed double dotted lines represent the results for starch B with $\left[\overline{M}_n, \overline{M}_w, N, PD\right] = [4100, 5430, 22496, 1.324]$, while solid lines represent the results for starch C with $\left[\overline{M}_n, \overline{M}_w, N, PD\right] = [160, 212, 878, 1.325]$ 169

LIST OF TABLES

Table 4.1: Values of parameters (Breuninger et al., 2009) used in the case study on chain-end scission	55
Table 5.1: The components of \mathcal{S} and E for one possible scenario of enzyme complex formation between two substrates of different chain length and two different enzymes. The number of enzymes and substrates are generally not restricted to two.	106
Table 5.2: Values of model parameters used in case study I where the specified initial ranges for calibration using the Genetic Algorithm (GA) were deduced by bracketing the extreme values reported by several similar studies in the literature (Altintas et al., 2002; Gadgil et al., 1996; Jang & Chou, 2013; Kobayashi & Nakamura, 2003, 2004; Ochoa et al., 2007). Calibration was done using the data reported by Nakamura et al. (1997)...	123
Table 5.3: Initial conditions used in case study I. For the population balance component, the symbol $m_S(0)$ is the initial mass concentration of starch, \overline{M}_n is the number-average DP, and \overline{M}_w is the weight-average DP.	126
Table 5.4: Values of model parameters used in case study II where the specified initial ranges for calibration using the Genetic Algorithm (GA) were deduced by bracketing the extreme values reported by several similar studies in the literature (Altintas et al., 2002; Birol, Önsan, Kırdar, & Oliver, 1998; Gadgil et al., 1996; Jang & Chou, 2013; Kobayashi & Nakamura, 2003, 2004; Ochoa et al., 2007). Calibration was done using the data reported by Ülgen et al. (2002).	149
Table 5.5: Initial conditions used in case study II. For the population balance component, the symbol $m_S(0)$ is the initial mass concentration of starch, \overline{M}_n is the number-average DP, and \overline{M}_w is the weight-average DP.	153

Table 5.6: Key assumptions that can be relaxed without compromising the interlinked PBM and CM framework.....	171
--	-----

LIST OF AND SYMBOLS AND ABBREVIATIONS

A	:	Mass concentration of ethanol (g/L)
A_n^α	:	Subsite affinity of the n -th subsite of α -amylase (kcal/mol)
A_n^γ	:	Subsite affinity of the n -th subsite of glucoamylase (kcal/mol)
B	:	Biomass
B'	:	Biomass excluding the key enzyme
b_{ij}^α	:	Discrete stoichiometric kernel for random scission relating the formation of polymers represented by the i -th pivot from polymers at the j -th pivot
$b(v,w)$:	Continuous stoichiometric kernel relating the formation of polymers with DP = v and $w-v$ from w
$b^\alpha(v,w)$:	Continuous stoichiometric kernel for random scission
$b^\gamma(w-v_m, w)$:	Continuous stoichiometric kernel for chain-end scission
CM	:	Cybernetic Modelling
C_g	:	Concentration of gelatinized starch
C_i	:	Molar concentration of the i -th interval (mol/L)
$C_{B,i}^\alpha$:	Molar concentration of sugar i bounded by α -amylase (mol/L)
$C_{B,i}^\gamma$:	Molar concentration of sugar i bounded by glucoamylase (mol/L)
$C^{(1)}$:	First moment of a polymer distribution
$C^{Exact}(t_j)$:	Exact molar concentration at time t_j
$C^{FP}(t_j)$:	FP molar concentration at time t_j

$\mathbf{C}^{(w)}(t_k)$:	Partitioned molar concentration vector evaluated at discrete time steps
c_i	:	Discrete molar concentration density of the i -th interval
$c(v,t)$:	Continuous molar concentration density
DNS	:	3,5-dinitrosalicylic acid
DP	:	Degree of Polymerization
\mathbf{E}	:	Molar concentration vector of all the free (or unbounded) enzymes
$\mathbf{E}(0)$:	Vector of initial enzyme loading
\mathbf{E}^0	:	Total molar concentrations of extracellular enzymes at a particular instance
E_α^0	:	Total molar concentration of α -amylase (mol/L)
E_γ^0	:	Total molar concentration of glucoamylase (mol/L)
E_α	:	Concentration of unbounded enzyme exhibiting random scission (α -amylase)
E_γ	:	Concentration of unbounded enzyme exhibiting chain-end scission (glucoamylase)
\mathbf{e}	:	Vector of enzyme levels with the size of n_e in the CM
e_i	:	Mass concentration of key metabolic enzyme (g/L)
$e_{\max,i}$:	Maximum level of e_i (g/g-DW)
$e_i/e_{\max,i}$:	Relative level of enzyme i
FP	:	Fixed Pivot
GA	:	Genetic Algorithm
H	:	Total number of time steps in a simulation
J_{opt}	:	Objective function for nonlinear parameter estimation

K_d	:	Rate constant for cellular death (1/h)
K_{ETOH}	:	Constant of ethanol inhibition on growth (g/L)
$K_{E,\alpha}$:	Kinetic constants of extracellular α -amylase synthesis (g/L)
$K_{E,\gamma}$:	Kinetic constants of extracellular glucoamylase synthesis (g/L)
$K_{e,i}$:	Kinetic constant for intracellular enzyme synthesis (g/L)
K_i	:	Saturation constant of sugar i (g/L)
K_α	:	Saturation constant for r_α (g/L)
K_γ	:	Saturation constant for r_γ (g/L)
$K_{J,i}^\alpha$:	Association constants of the i -mer substrate in a binding mode J by α -amylase (which may either be productive or non-productive) (mol/L)
$K_{J,i}^\gamma$:	Association constants of the i -mer substrate in a binding mode J by glucoamylase (which may either be productive or non-productive) (mol/L)
$K_{m,i}^\alpha$:	M-M parameter for hydrolysis by α -amylase (mol/L)
$K_{m,i}^\gamma$:	M-M parameter for hydrolysis by glucoamylase (mol/L)
K_I^α	:	Constant of α -amylase repression by glucose (g/L)
K_I^γ	:	Constant of glucoamylase repression by glucose (g/L)
\mathbf{k}	:	Vector of rate constants for the PBM component
$k_{E,\alpha}$:	Kinetic constants of extracellular α -amylase synthesis (1/h)
$k_{E,\gamma}$:	Kinetic constants of extracellular glucoamylase synthesis (1/h)
$k_{e,i}$:	Kinetic constant for intracellular enzyme synthesis (1/h)
k_i^α	:	Discrete rate kernel for random scission

k_i^γ	:	Discrete rate kernel for chain-end scission
$k(v)$:	Continuous rate kernel
$k^\alpha(v)$:	Continuous rate kernel for random scission
$k^\gamma(v)$:	Continuous rate kernel for chain-end scission
$k_{a,i}^\alpha$:	Rate constant of polymer with DP = i binding with α -amylase (L/mol/h)
$k_{a,i}^\gamma$:	Rate constant of polymer with DP = i binding with glucoamylase (L/mol/h)
$k_{b,i}^\alpha$:	Rate constant of polymer with DP = i detaching from α -amylase (1/h)
$k_{b,i}^\gamma$:	Rate constant of polymer with DP = i detaching from glucoamylase (1/h)
$k_{c,i}^\alpha$:	Rate constant of hydrolysis of polymer with DP = i by α -amylase (1/h)
$k_{c,i}^\gamma$:	Rate constant of hydrolysis of polymer with DP = i by glucoamylase (1/h)
M-M	:	Michaelis-Menten
MWD	:	Molecular Weight Distribution
M_i	:	Molecular weight of the i -th mer
M_α	:	Molecular weight of α -amylase (g/mol)
M_γ	:	Molecular weight of glucoamylase (g/mol)
\overline{M}_n	:	Number-average DP
\overline{M}_w	:	Weight-average DP
$m_i(0)$:	Initial mass concentration of DP = i (g/L)
$m_s(0)$:	Initial mass concentration (mass/volume) of the polymer

m_1^{eq}	:	Total mass of substrate in the form of glucose equivalent (g)
N	:	Maximum DP
\mathcal{N}_i	:	Mass concentration of substrates in the abiotic phase (g/L)
N_m^α	:	Total number of subsites of the α -amylase
N_m^γ	:	Total number of subsites of the glucoamylase
n_{ij}	:	Fractional allocation of polymers splitting from DP = j into i
n_{ij}^α	:	Fractional allocation of polymers splitting from DP = j into i due to random scission
n_{ij}^γ	:	Fractional allocation of polymers splitting from DP = j into i due to chain-end scission
ODEs	:	Ordinary Differential Equations
PBEs	:	Population Balance Equations
PBM	:	Population Balance Modelling
PDE	:	Partial Differential Equation
PD	:	Polydispersity Index
\mathcal{P}	:	Mass concentration of extracellular products (g/L)
$P(i)$:	Polymer with DP = i
p	:	Number of pivots in the discrete region
p_{min}	:	Minimum value of p
q	:	Number of pivots in the continuous region
RHS	:	Right Hand Side
RS	:	Reducing sugar (g/L)
R	:	Universal gas constant (kcal/K/mol)
R_k	:	Return on investment from the k -th alternative

\mathbf{r}	:	Regulated fluxes (or rates defined per unit of biomass)
\mathbf{r}_e	:	Inducible enzyme synthesis rates
r	:	Geometric ratio
$r_{E,\alpha}$:	Rate of α -amylase synthesis (g/g-DW/h)
$r_{E,\gamma}$:	Rate of glucoamylase synthesis (g/g-DW/h)
$r_{X,i}$:	Specific rate of biomass growth on sugar i (g/g-DW/h)
r_α	:	Specific rate of indirect biomass growth on starch through α -amylase (g/g-DW/h)
r_γ	:	Specific rate of indirect biomass growth on starch through glucoamylase (g/g-DW/h)
SSF	:	Simultaneous Saccharification and Fermentation
\mathcal{S}	:	Size of an individual partition of the molar concentration vector
\mathcal{S}	:	Vector of all substrate-related terms (including the complexes)
$\mathcal{S}^{(\xi)}$:	ξ -th moment of the polymer population \mathcal{S}
T	:	Temperature
t	:	Time
t_p	:	Time required to achieve peak ethanol concentration
$t_{99\%}$:	Time required for the molar concentration of the monomer to reach 99% of its final value
t_{Exact}	:	Time required to obtain the exact solution
t_{FP}	:	Time required to obtain the fixed pivot solution
U_e	:	Cybernetic variable for regulating the synthesis of intracellular enzymes
U_E	:	Cybernetic variable for regulating the synthesis of extracellular enzymes

$U_{E,\alpha}$:	Cybernetic variable for regulating the synthesis of α -amylase
$U_{E,\gamma}$:	Cybernetic variable for regulating the synthesis of glucoamylase
$U_{e,i}$:	Cybernetic variable for regulating the synthesis of intracellular enzyme i
U_i	:	Cybernetic variable for regulating the synthesis of enzyme i
V_i	:	Cybernetic variable for regulating the activity of enzyme i
v	:	Continuous degree of polymerization
v_m	:	DP of the monomer
$W_{\mathcal{N}}$:	Matrix containing the stoichiometry of substrate consumption in the CM
$W_{\mathcal{P}}$:	Matrix containing the stoichiometry of product formation in the CM
W_{ϕ}	:	Matrix containing the stoichiometry of intracellular metabolite accumulation in the CM
\bar{W}_i	:	Weights assigned to different variables in J_{opt}
w	:	Total number of partitions for the molar concentration vector
X	:	Mass concentration of biomass (g-DW/L)
x_i	:	DP of the i -th pivot for the FP solution
Y_i	:	Biomass yield from the utilization of sugar i (g/g)
Y_i^{EtOH}	:	Ethanol yield from the utilization of sugar i (g/g)
y_i	:	DP of the i -th pivot for the exact solution
\hat{y}_i	:	Vector of predicted output
\tilde{y}_i	:	Vector of experimental observations
Z	:	Upper triangular coefficient matrix

$\Gamma(\omega)$:	Gamma function
Θ_i	:	i -th parameter to be calibrated
$\Theta_{i,\min}$:	Lower bound for Θ_i
$\Theta_{i,\max}$:	Upper bound for Θ_i
β	:	Enzyme degradation rate constants
β_{EtOH}	:	Deactivation of glucoamylase within the broth ($L^n/g^n/h$)
$\beta_{E,\alpha}$:	Deactivation rate constant for α -amylase (1/h)
$\beta_{E,\gamma}$:	Deactivation rate constant for glucoamylase (1/h)
$\beta_{e,i}$:	Intracellular enzyme degradation rate constant (1/h)
ε_g	:	Global error indicator to quantify deviations from the exact solution
ε_S	:	Error in the initial mass of polymers due to discretization
η	:	Order of ethanol inhibition on glucoamylase production
θ	:	Dimensionless time
λ_i	:	Vector of eigenvalues for $\mathbf{G}^{(i,i)}$
μ	:	Total specific growth rate
$\mu_{\max,i}$:	Maximum specific growth rate of sugar i (1/h)
μ_α	:	Specific growth rate for r_α (1/h)
μ_γ	:	Specific growth rate for r_γ (1/h)
ρ_e	:	Constitutive intracellular enzyme synthesis rates
ρ_E	:	Constitutive extracellular enzyme synthesis rates
τ_i	:	Number of observations for the variable i
ϕ	:	Concentration of intracellular metabolites

LIST OF APPENDICES

APPENDIX A: MODELLING CHAIN-END SCISSION	196
A.1 Theoretical Preliminaries of The Fixed Pivot Technique (S. Kumar and Ramkrishna, 1996a).....	196
A.2 MATLAB Code for The Simulation of Chain-End Scission using The Fixed Pivot Technique.....	200
A.3 MATLAB Code for The Simulation of Chain-End Scission using The Exact Solution	203
A.4 MATLAB Code for The Simulation of Chain-End Scission on Stickel's Example (Stickel and Griggs,2012) using The Fixed Pivot Technique	205
A.5 MATLAB Code for The Simulation of Chain-End Scission on Stickel's Example (Stickel and Griggs, 2012) using The Exact Solution.....	207
A.6 MATLAB Code for Calculating p_{min}	209
APPENDIX B: MODELLING RANDOM SCISSION.....	210
B.1 Deriving The Fixed Pivot Equations for Discrete Random Scission.....	210
B.2 Simplifying The Fixed Pivot Expression for Random Scission	213
B.3 Proof that the Fixed Pivot Technique Over-Predicts for Random Scission..	214
B.4 MATLAB Code for The Simulation of Random Scission on Stickel's Example (Stickel and Griggs, 2012) Using The Fixed Pivot Technique	217
B.5 MATLAB Code for The Simulation of Random Scission on Stickel's Example (Stickel and Griggs, 2012) Using The Exact Solution	220
APPENDIX C: MODELLING SIMULTANEOUS RANDOM AND CHAIN-END SCISSION.....	222

C.1	MATLAB Code for The Simulation of Simultaneous Random and Chain-End Scissions Using The Fixed Pivot Technique	222
C.2	MATLAB Code for The Simulation of Simultaneous Random and Chain-End Scissions Using The Exact Solution.....	226
APPENDIX D: EXAMPLES OF PBM IMPLEMENTATION ON ENZYMATIC HYDROLYSIS		229
APPENDIX E: KINETICS OF MICROBIAL FERMENTATION		231
E.1	Examples of Kompala's Cybernetic Model Written in The General Form ..	231
E.2	Rationale for Not Equating $k_{e,i} = \mu_{\max,i} + \beta_{e,i}$	233
APPENDIX F: EXTRACELLULAR DEPOLYMERASE PRODUCTION		235
F.1	Derivation of The Extracellular Enzyme Production Equation	235
F.2	Formulation for r_j ($j = \gamma$ or α).....	236
APPENDIX G: SUBSITE THEORY		238
G.1	Subsite Theory for Glucoamylase.....	238
G.2	Subsite Theory for α -amylase	240
G.3	Subsite Affinity Maps	242
APPENDIX H: MATLAB CODE FOR THE SIMULATION OF THE SSF OF A GLUCOAMYLASE PRODUCING RECOMBINANT YEAST		245
APPENDIX I: MATLAB CODE FOR THE SIMULATION OF THE SSF OF AN α -AMYLASE AND GLUCOAMYLASE PRODUCING RECOMBINANT YEAST ...		252

CHAPTER 1 : INTRODUCTION

1.1 Research Background

Natural polymers, e.g. starch, cellulose etc., are important raw materials in the generation of valuable products such as fuel ethanol (Jang & Chou, 2013; G. S. Murthy, Johnston, Rauseh, Tumbleson, & Singh, 2011) and beer (Brandam, Meyer, Proth, Strehaiano, & Pingaud, 2003; Koljonen, Hämäläinen, Sjöholm, & Pietilä, 1995; Marc & Engasser, 1983) through fermentation. During fermentation, these macromolecules cannot be directly consumed by the microbes. Thus, they have to be first broken down to simple substrates to facilitate subsequent assimilation by the microbes for conversion to useful products. Amongst many other methods, enzymatic hydrolysis is one most promising approach (El-Zawawy, Ibrahim, Abdel-Fattah, Soliman, & Mahmoud, 2011). While commercial enzyme preparations can be dosed in to the fermentation broth for this purpose (C. G. Lee, Kim, & Rhee, 1992; T. Montesinos & J-M. Navarro, 2000), it is increasingly common to have the microbes produce the enzymes *in situ* (Altintas, Kirdar, Onsan, & Ulgen, 2002; Azmi, Ngoh, Mel, & Hasan, 2010; Kroumov, Módenes, & de Araujo Tait, 2006; Ochoa, Yoo, Repke, Wozny, & Yang, 2007). In short, microorganisms are employed to produce the depolymerization enzymes as well as to convert the depolymerized components into valuable products. Under such circumstances, the biochemistry of the process is highly complex. Being polydisperse in nature, the Molecular Weight Distribution (MWD) of natural polymers subject to enzymatic depolymerization is constantly evolving, and this creates a highly complex broth with numerous substrates, of which only those within a certain size range can be assimilated by the microbes. Given multiple substrate choices, the course of fermentation will be affected by the response of the microbes toward each potential nutrient. Accurate

knowledge of such information is particularly important to the practitioners as they attempt to design more efficient processes.

In enzymatic depolymerization, the temporal evolution of the MWD varies according to the mode of action exhibited by the enzymes. Two commonly encountered enzymatic depolymerization phenomena are random and chain-end scission. For random scission, bond cleavages occur randomly along the bonds within a polymer chain while enzymes which exhibit chain-end scission behaviour remove a fixed number of mers from the end of a polymer chain. Examples of enzymes which exhibit random scission are α -amylase (EC 3.2.1.1) and endo-cellulases (EC 3.2.1.4) etc., whereas enzymes which exhibit chain-end scission are glucoamylase (EC 3.2.1.3), β -amylase (EC 3.2.1.2), and exo-cellulases (EC 3.2.1.91) etc. The traditional kinetic approach to model enzymatic scission is to employ the Michaelis-Menten (M-M) type of expressions by treating the natural polymer as a grossly lumped entity (Kusunoki, Kawakami, Shiraishi, Kato, & Kai, 1982; Miranda & Murado, 1991; Nakamura, Kobayashi, Ohnaga, & Sawada, 1997; Polakovič & Bryjak, 2004; Presečki, Blažević, & Vasić-Rački, 2013; Shiraishi, Kawakami, & Kusunoki, 1985). Such an approach is incapable of distinguishing between different modes of enzymatic action and does not track the transient of the entire MWD. To capture these details, the Population Balance Modelling (PBM) technique (D. Ramkrishna, 2000; Doraiswami Ramkrishna & Singh, 2014) is a suitable framework. In this technique, the temporal evolutions of all the polymer populations as well as the mode of enzymatic action are considered, thus allowing a more fundamental analysis of the depolymerization phenomena. The PBM technique was employed in several studies for the modelling of pure enzymatic scission of natural polymers, e.g. for the hydrolysis of cellulose (Griggs, Stickel, & Lischeske, 2012a, 2012b; Hosseini & Shah, 2011a, 2011b) and starch (Chang, Delwiche, & Wang, 2002).

The course of converting the products of enzymatic hydrolysis to useful products occurs mainly through the metabolism of microorganisms. In this process, substrates present in the abiotic phase are assimilated into the cells and go through a series of transformation reactions via complex metabolic networks with the production of intracellular metabolites as well as other metabolic products which eventually are excreted into the external environment of the cells. Cellular metabolism is also regulated internally in terms of the levels and activities of the key enzymes catalyzing specific reactions. Therefore, given a choice of different substrates, i.e. carbon and energy source, the rates at which these different substrates are assimilated into the biotic phase are determined by these in-built regulatory mechanisms. In the context of microbial conversion of starch hydrolysates, one example is the fermentation of sugars to ethanol by yeast where it is well established that in addition to glucose, maltose and maltotriose may also be fermented to produce ethanol, and the presence of glucose generally represses the consumption of these larger sugars (Duval, Alves Jr, Dunn, Sherlock, & Stambuk, 2010; Ernandes, D'Amore, Russell, & Stewart, 1992; T. Montesinos & J.-M. Navarro, 2000). In view of the capability of microbes in responding to diverse environmental changes, any effort in dynamic modelling must account for the effects of regulation. The Cybernetic Modelling (CM) approach, spearheaded by Ramkrishna and co-workers (cf. D. Ramkrishna and Song (2012) for a review of the developments) was thus developed in this light. Differing from the traditional approach in modelling physical systems which places a heavy demand for mechanistic information, the CM approach posits that metabolic regulation is attached to a certain goal-seeking behaviour of the organism. Using what they refer to as the “cybernetic variables”, the levels and activities of key enzymes necessary for affecting the microbial response to environmental changes are regulated by the minimization/maximization of an objective function. Evolving over the past three decades, this main feature of the CM approach has enabled it to successfully

predict the growth of various microbes in complex substrate environment --- from the simplest variant employed by Kompala and Ramkrishna (1986) to the more advanced variant implemented by Geng, Song, Yuan, and Ramkrishna (2012). For the case of fermenting the products of enzymatic hydrolysis, different variants of the cybernetic models, i.e. those developed by Kompala and Ramkrishna (1986) as well as by Varner and Ramkrishna (1998), were employed as part of the Simultaneous Saccharification and Fermentation (SSF) framework for both starch (Altintas et al., 2002; Chavan, Raghunathan, & Venkatesh, 2009; Ganti S. Murthy, Johnston, Rausch, Tumbleson, & Singh, 2012; Ochoa et al., 2007) as well cellulose (Ko et al., 2010) and lignocellulose (Shin, Yoo, Kim, & Yang, 2006) conversions.

Insofar as the simultaneous scission and conversion of natural polymers is concerned, the most common modelling effort employed to date has been that of coupling the lumped parameter M-M type models with the unstructured models for microbial growth (Anuradha, Suresh, & Venkatesh, 1999; Hofvendahl, Åkerberg, Zacchi, & Hahn-Hägerdal, 1999; Jang & Chou, 2013; C. G. Lee et al., 1992; Morales-Rodriguez, Gernaey, Meyer, & Sin, 2011; Ochoa et al., 2007; Podkaminer, Shao, Hogsett, & Lynd, 2011). While the M-M type models are a gross over-simplification of the depolymerization phenomena, the use of the unstructured models for microbial growth kinetics lacks the necessary robustness to handle complex nutrient environments (D. Ramkrishna & Song, 2012). Employment of the CM approach improves prediction with regards to the microbial kinetics, but for SSF it has thus far only been primarily coupled to the M-M type models for enzymatic scission (Altintas et al., 2002; Chavan et al., 2009; Ko et al., 2010; Shin et al., 2006), thereby ignoring the details of enzymatic scission at the molecular level. As the enzymatic scission and the microbial growth are both major components of the SSF process, successful mathematical abstraction of the process is

therefore closely dependent on the level of essential details given to each component. For this purpose, the PBM and the CM frameworks as alluded to previously are known to be excellent in capturing the critical details of the respective individual component (D. Ramkrishna & Song, 2012; Sterling & McCoy, 2001). Despite being the method par excellence in their own areas of application, the mathematical linkage between the two conceptually different techniques for the abstraction of the SSF process appears not to have been established. One possible reason might be the distinct mathematical nature of the two: the CM is formulated as a system of Ordinary Differential Equations (ODEs) which can be readily integrated with commercial solvers while the PBM involves the solution of a Partial Differential Equation (PDE) with an integral term, thus requiring the use of special techniques (M. Kostoglou, 2007; M. Kostoglou & Karabelas, 2002, 2004, 2009; J. Kumar, Peglow, Warnecke, & Heinrich, 2008; S. Kumar & Ramkrishna, 1996a, 1996b; D. Ramkrishna, 2000). Another plausible reason is the lack of cross-talk due to disciplinary differences: users of CM are more from the biochemical engineering fraternity (Shin et al., 2006; Song, Morgan, & Ramkrishna, 2009), while users of PBM are more in particulate technologies (Atmuri, Henson, & Bhatia, 2013; J. Kumar et al., 2008; Nopens, Beheydt, & Vanrolleghem, 2005) and synthetic polymer (Alexopoulos, Pladis, & Kiparissides, 2013; Staggs, 2005).

In this work, the general linkage between the two conceptually different approaches, i.e. the PBM and the CM, was for the first time established and introduced in a single framework. Figure 1.1 illustrates the general idea. The two critical links are denoted by the hedged arrows. The first involves the withdrawal of the small oligomers produced by enzymatic breakdown of large polymer chains by the microbes. The second involves the excretion of extracellular depolymerization enzymes into the fermentation broth for the breaking down of polymeric substrates. These links enable interactions between both

processes as described by PBM and CM. To PBM, instead of the classic accumulation of smaller sugars in the solution, their consumption might occur. In addition, the enzyme levels responsible for depolymerization might change depending on the output from CM. Both effects have to be incorporated into PBM. Likewise, to CM, the substrate concentrations will be much more dynamic instead of being merely depleting over time. Conceivably the microbes may have to switch back and forth between various preferred substrates as their relative abundance evolves over time. More critically, CM has to include the excretion of extracellular depolymerase, i.e. enzymes for depolymerization. If only large polymers are present, the depolymerase should be induced. When preferred substrates are formed, the depolymerase excretion should eventually be repressed. However, the previously released depolymerase remains active for a finite duration in the broth, and this must be recognizable by CM.

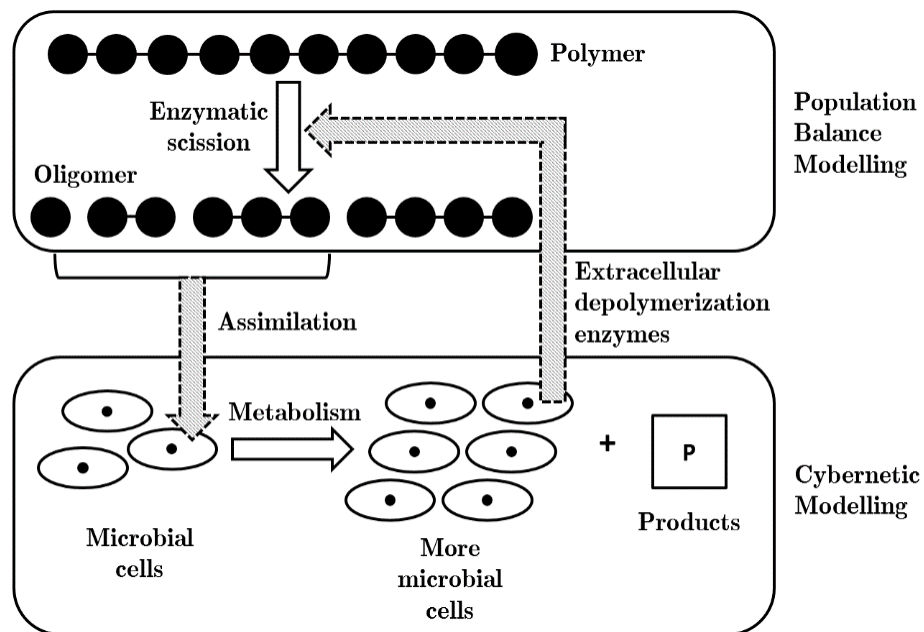


Figure 1.1: Illustrating the linkage between the PBM and CM in a single framework. In particular, the induction/repression of the synthesis of extracellular depolymerase by the composition of the broth appears to be relatively unexplored within the CM framework.

In the work of Gadgil, Bhat, and Venkatesh (1996), the CM framework was used to model the excretion of extracellular α -galactosidase for the breakdown of disaccharide melibiose to glucose and galactose, of which the presence of glucose represses the assimilation of galactose and the excretion of α -galactosidase. The model involved a simplistic assumption that the concentration of α -galactosidase corresponded to the concentration of the key enzyme involved in metabolizing galactose, since both are repressed by glucose. This assumption was later also employed by Altintas et al. (2002) for the excretion of extracellular fusion protein displaying both α -amylase and glucoamylase activities for the breakdown of starch to glucose and reducing sugars. In their case, the presence of glucose represses the assimilation of reducing sugars, and thus the depolymerase was assumed to correspond to the key enzyme for metabolizing the reducing sugars. This strategy is only applicable when only two substrates are capable of being metabolized. For the breakdown of large polymers such as starch, clearly the microbes are incapable of consuming every reducing sugar in the broth but that only the smaller ones are consumed, e.g. glucose, maltose, and maltotriose for the yeast *Saccharomyces cerevisiae* (Duval et al., 2010; Ernandes et al., 1992; T. Montesinos & J.-M. Navarro, 2000). As such, if the microbes are capable of metabolizing more than two substrates resulting from the breakdown of polymers, the choice of the key enzymes used to mimic the excretion of extracellular depolymerase is not apparent. Another contribution of this work is to clarify this choice in a systematic manner.

1.2 Research Objectives

The main objectives of this research are as follow:

- a) To select or modify a potential numerical technique for approximating and solving population balance equations for both chain-end and random scissions as well as their combination thereof. This includes exploration of the inherent characteristics of the resulting formulation.
- b) To interlink the PBM and CM for modelling the batch growth of a microbial strain capable of simultaneously hydrolyzing a natural polymer and fermenting the resulting smaller saccharides. The resulting model will be used to analyze the growth of microbes on complex nutrients resulting from the individual or the combined actions of enzymes exhibiting random and chain-end scission behaviour.

1.3 Structure of the Thesis

The remaining six chapters are organized as follows:

- a) Chapter 2 reviews the pertinent literature of this research. A brief introduction to SSF is given followed by a review of the common methodologies employed in the modelling of SSF processes. After that, the PBM is introduced with specific reference to random and chain-end scissions, highlighting the state of solution techniques which are necessary and appropriate for achieving the objectives of this work. Moreover, the use of the PBM in modelling saccharification processes is also

reviewed. Following this, the mathematical background for the CM framework is given and past employments of the CM framework in modelling SSF are elaborated.

- b) Chapter 3 discusses the general methods used in meeting all the research objectives. The theoretical formulations and detailed mathematical derivations are deferred to Chapters 4 and 5 preceding the presentation of the results and discussions.
- c) Chapter 4 is the fulfilment of the first objective of this work. In this chapter, the theoretical formulations for solving population balance equations involving both chain-end and random scissions using the fixed pivot technique are deliberated. Upon establishing the necessary solution technique, the results are benchmarked against the exact solution. Further observations of the fixed pivot technique in solving chain-end and random scissions are also presented.
- d) Chapter 5 fulfils the second objective where the general framework for interlinking the PBM and the CM components are presented. This is followed by a demonstration of the capability of the resulting framework in modelling the SSF processes. Two case studies are used for this purpose, the first involves the growth on starch of a yeast producing glucoamylase (chain-end scission enzyme) and the second for the growth on starch of a yeast producing both glucoamylase and α -amylase (random scission enzyme).
- e) Chapter 6 summarizes the key findings of this work, and proposes future extensions.

CHAPTER 2 : LITERATURE REVIEW

2.1 Simultaneous Saccharification and Fermentation (SSF)

The use of biological systems, i.e. microorganisms and biocatalysts, to produce valuable products is an ancient idea, e.g. in the making of cheese (Caplice & Fitzgerald, 1999), wine (Pretorius, 2000) etc. Nevertheless, the technologies of bioprocessing are constantly evolving in the direction which enable the processes to be operated more efficiently and at the same time meeting specific economic objectives. One such example is the production of ethanol which finds its usefulness in the many facets of life, one of which is an alternative to fuel (G. S. Murthy et al., 2011). To produce ethanol, a well established biochemical route is to metabolize sugar with a selected microbial strain, e.g. *S. cerevisiae* under anaerobic condition (Shuler & Kargi, 2002). However, the search for more renewable feedstock has led to the use of polymeric substrates such as starchy, cellulosic and lignocellulosic biomass (Bansal, Hall, Realff, Lee, & Bommarius, 2009; Jang & Chou, 2013; van Maris et al., 2006) which must be hydrolyzed to yield the required substrates for subsequent microbial fermentation. As such, the two important stages in the production of ethanol (excluding the pre-treatment of raw materials) are the breaking of polymeric bonds through enzymatic hydrolysis followed by the conversion of smaller sugars through fermentation. Although these two steps are traditionally carried out separately because of the differences in the optimal operating conditions, it is an increasingly popular practice to consolidate the two stages by performing both hydrolysis and fermentation simultaneously (Taherzadeh & Karimi, 2007) – hence, the term ‘Simultaneous Saccharification and Fermentation (SSF)’. Other than being employed in the production of ethanol, SSF is also being employed for the production of other chemicals such as lactic acid (R. P. John, Nampoothiri, & Pandey, 2009) and fumaric

acid (Deng, Li, Xu, Gao, & Huang, 2012) by using different strains. During SSF, the enzymes required for breaking the polymeric bonds can either be dosed externally or it can be biologically produced *in situ*. In the latter case, if there is no single strain capable of producing the enzymes and fermenting the substrates simultaneously, a concoction of different strains serving different purpose may be employed, e.g. as was done in the study of Azmi et al. (2010). One of the benefits of SSF over the separate hydrolysis and fermentation process is its ability to minimize the effect of inhibition (on the enzymes or the microbes) due to the continuous withdrawal of sugars from the fermentation broth through microbial consumption (Taherzadeh & Karimi, 2007). This contributed to higher product yield in the case of ethanol production compared to the traditional route (Eklund & Zacchi, 1995; Karimi, Emtiazi, & Taherzadeh, 2006; McMillan, Newman, Templeton, & Mohagheghi, 1999; Sun & Cheng, 2002). Depending on the implementational configurations, i.e. the origin of the enzymes used (whether dosed externally or produced by the microbes) and the number of strains involved, the SSF process can be very complicated to model and reported attempts to do this will be reviewed next.

2.1.1 Modelling of SSF Processes

The modelling of any process involving biological systems is non-trivial. This is more so true in the case where both enzymatic scissions and fermentation are carried out concurrently. Modelling of such processes requires the integration of two essential components, viz. the model for polymeric scissions as well as the model for microbial growth. The most common framework used to date is to employ the M-M type model for the enzymatic scission of polymer molecules and the Monod type kinetics for cell growth. As the Monod type kinetics for cell growth are unstructured and do not account for the intracellular components within the cell, the forms of expression used do not vary much

from those reported in standard biochemical engineering textbook, e.g. Shuler and Kargi (2002). In this regard, the differences in the expressions used are mainly due to assumptions on inhibition of cell growth by different components of the fermentation broth (e.g. ethanol, glucose etc.), the inclusion of maintenance considerations (Nakasaki, Murai, & Akiyama, 1988; Shen & Agblevor, 2011) or cellular death (Ochoa et al., 2007).

As opposed to the model expressions for cell growth which are rather straightforward, the expressions for enzymatic scission of polymeric substrates vary according to the assumptions made with regard to the mechanism of scission. As the M-M type kinetic models are derived on the basis of the interaction between the enzyme and the substrate molecules, different expressions can be obtained with different mechanistic assumptions. One such example is the possibility of inhibition by different molecules on the enzyme activity. Such inhibitions may be initiated by the products (Anuradha et al., 1999; Hofvendahl et al., 1999; C. G. Lee et al., 1992; Morales-Rodriguez, Meyer, Gernaey, & Sin, 2011; Nakasaki et al., 1988; Podkaminer et al., 2011; Shen & Agblevor, 2010; Shen & Agblevor, 2011; van Zyl, van Rensburg, van Zyl, Harms, & Lynd, 2011; Zhang, Shao, Townsend, & Lynd, 2009), the substrates (Jang & Chou, 2013) or both (Davis, 2008; Kroumov et al., 2006; Philippidis, Spindler, & Wyman, 1992). Other than that, as the substrates involved in the SSF processes are polymers, the scission of these molecules produces many different products of different chain lengths. Depending on which intermediate products are to be accounted for in the model as well as the network of reactions assumed, the resulting systems of M-M type equations vary from one network to another. In the simplest form of reaction network, the initial substrate is assumed to be depolymerized to the final product without the production of intermediates, e.g. from cellulose to glucose (Nakasaki et al., 1988), starch to glucose (Anuradha et al., 1999; Jang & Chou, 2013; Kroumov et al., 2006; Matsumura & Hirata, 1989) etc. This is, of course,

a gross over-simplification of polymeric scission. Other authors expanded the reaction network by including more intermediate products such as the hydrolysis of cellulose following the order of cellulose \rightarrow cellobiose \rightarrow glucose (Morales-Rodriguez, Gernaey, et al., 2011; Morales-Rodriguez, Meyer, et al., 2011; Philippidis et al., 1992; Podkaminer et al., 2011; Shen & Agblevor, 2010; Shen & Agblevor, 2011; van Zyl et al., 2011; Zhang et al., 2009) or the hydrolysis of starch to various smaller oligomers (Davis, 2008; Hetényi, Németh, & Sevela, 2011; Hofvendahl et al., 1999; C. G. Lee et al., 1992). Such efforts although mathematically more developed are nonetheless still a departure from the true nature of polymeric scission.

The type of enzymes employed with their respective mode of action also plays a part in dictating the model expressions. Two different types of enzymatic actions are typically encountered, i.e. random scission and chain-end scission. Although these are clearly distinct mode of actions, their mechanistic pronouncement in the M-M type models within the SSF framework is still unclear. For instance, various authors made no distinction between the actions of endo- and exo-acting enzymes, thus treating them as a single entity (Davis, 2008; Hetényi et al., 2011; Morales-Rodriguez, Gernaey, et al., 2011; Morales-Rodriguez, Meyer, et al., 2011; Nakasaki et al., 1988; Philippidis et al., 1992; Podkaminer et al., 2011; Shen & Agblevor, 2010; Shen & Agblevor, 2011; Zhang et al., 2009). An attempt was made by van Zyl et al. (2011) to separate the action of endo- and exo-acting enzymes. However, certain simplifying assumptions which were made did not reflect the depolymerization phenomena mechanistically where the cellulose chains in the presence of both endo- and exo-acting enzymes were assumed to be converted to cellobiose by the sole action of the latter. Anuradha et al. (1999), in the attempt to model the action of exo-acting enzyme (i.e. glucoamylase) on starch, pre-incubated the substrate with endo-acting enzyme (i.e. α -amylase) but did not include the kinetics of the latter in

the model. As such, these models even though adopting the M-M framework, are at best semi-empirical when faced with the complexity of polymeric scission.

Other than the above, consideration of the different structure of substrates will also affect the resulting model equations. Such heterogeneity is not uncommon where for instance gelatinized starch is composed of linear and branched components while cellulose can be divided into crystalline and amorphous regions. To simplify modelling efforts, heterogeneity of substrates was ignored in many studies (Anuradha et al., 1999; Davis, 2008; Hetényi et al., 2011; Hofvendahl et al., 1999; Jang & Chou, 2013; Matsumura & Hirata, 1989; Morales-Rodriguez, Gernaey, et al., 2011; Morales-Rodriguez, Meyer, et al., 2011; Nakasaki et al., 1988; Podkaminer et al., 2011; Shen & Agblevor, 2010; Shen & Agblevor, 2011; Zhang et al., 2009). To include the heterogeneity of substrates as part of the model, efforts to do this were done by structuring starch into linear vs. branched components (C. G. Lee et al., 1992) or resistant vs. susceptible portions (Kroumov et al., 2006; Ochoa et al., 2007). For cellulose this was accounted for by structuring cellulose into crystalline vs. amorphous regions (van Zyl et al., 2011).

The above summarizes the efforts made in modelling the SSF by integrating the M-M type models and the unstructured Monod type kinetics. This framework assumes a simplistic response from the microbes towards environmental changes by ignoring the intracellular components. To give a more realistic representation of microbial growth in SSF cultures, another framework had also been reported in the literature where the M-M type models were integrated with the chemically structured model (Shuler & Kargi, 2002) for microbial growth which captures the kinetic interaction between different cellular sub-components. In the work of Kobayashi and Nakamura (2003), a structured model for a

recombinant yeast culture was integrated with the M-M type model for the hydrolysis of starch by glucoamylase. The enzyme glucoamylase produced by the yeast culture was assumed to be induced by starch and repressed by the presence of glucose. This model was then adopted in the work of Kobayashi and Nakamura (2004) for an immobilized recombinant yeast culture. In another work, compartmental modelling was employed where the microbial cells were structured into three, four and eight compartments, each compartments containing cell components having similar functions (Arga et al., 2004). This form of structured model requires a greater level of mechanistic information which may be challenging to obtain.

Another form of structured model for microbial growth kinetics which had been integrated with the M-M type model is the cybernetic model developed by Ramkrishna and co-workers (D. Ramkrishna & Song, 2012). The approach is based on the argument that microbes are inherently able to regulate their own metabolism in fulfilment of a certain goal or objective. In the work of Altintas et al. (2002), the version of cybernetic model developed by Kompala and Ramkrishna (1986) was used to model the SSF of starch by a recombinant yeast which secretes a fusion protein displaying both α -amylase and glucoamylase activities. The same process was modelled by Ochoa et al. (2007) using the version of cybernetic model proposed by Varner and Ramkrishna (1999b) which incorporated more complex pathways. Other authors had also employed the CM framework to model the SSF of various starchy and lignocellulosic materials (Chavan et al., 2009; Ko et al., 2010; Ganti S. Murthy et al., 2012; Shin et al., 2006).

As evident from the discussion above, two components are essential to describing the SSF processes involving natural polymers. The first is the model for polymer scission which to date was mainly described by M-M type models. As the development of M-M

type models relies on the proposed mechanism of polymer scission, it is impractical to outline the mechanism of scission for the entire spectrum of molecules with various chain lengths. As such, the models employed are not infrequently an over-simplification of reality and at best semi-empirical in nature. Although more rigorous models such as those founded upon the subsite theory (Besselink, Baks, Janssen, & Boom, 2008; Marchal, Ulijn, de Gooijer, Franke, & Tramper, 2003; Marchal, Zondervan, Bergsma, Beeftink, & Tramper, 2001; Wojciechowski, Koziol, & Noworyta, 2001) or the population balance technique (Chang et al., 2002; Griggs et al., 2012a, 2012b) had been reported in the literature for enzymatic scission, they have yet to be assimilated into the SSF literature. The latter is a powerful approach for describing the evolution of polymer distributions which will be used in this work to model polymer scission. As for the kinetics of fermentation, the complex substrate environment for the fermentation broth requires that the model be sufficiently robust to capture the response of the microbes to changing substrate environment. As such, the structured model has the advantage over the unstructured model for being able to account for these changes through incorporation of metabolic details. In the face of inadequate mechanistic details on regulatory processes, the CM approach is a potentially powerful framework to model microbial kinetics. The PBM and CM approaches will be elaborated further in the following sections.

2.2 Population Balance Modelling (PBM)

As a polymer solution contains a large number of polymers of different sizes which cannot be studied individually via the conventional chemical reaction kinetics, the PBM technique is the natural approach as it considers the temporal evolution of the polymer distribution as part of solution. Assuming that the molar concentration density $c(v,t)$ is a continuous function of the degree of polymerization (v) and time (t), the general

continuous population balance model for binary scission is given as (D. Ramkrishna, 2000):

$$\frac{\partial c(v, t)}{\partial t} = 2 \int_v^\infty k(w) b(v, w) c(w, t) dw - k(v) c(v, t) \quad (2.1)$$

Here, $k(v)$ is the rate kernel which describes the rate of breaking a polymer with Degree of Polymerization, $DP = v$, and $b(v, w)$ is the stoichiometric kernel relating the formation of polymers with DPs of v and $w-v$ from w (McCoy & Wang, 1994). The factor “2” accounts for binary scission. The first term on the Right Hand Side (RHS) represents the scission of polymers with $DP \geq v$ to form v whereas the second term on the RHS represents the scission of polymers with $DP = v$ to form polymers with $DP < v$. The stoichiometric kernel is defined to satisfy the following constraints following the conservation of mass:

$$\int_0^w b(v, w) dv = 1 \quad (2.2)$$

$$b(v, w) = b(w - v, w) \quad (2.3)$$

Description of depolymerization phenomena using Eq. (2.1) is known as PBM. This equation can be solved to obtain various useful quantities such as the molar concentration of polymers with various chain lengths, the complete polymer distribution as well as the moments of the polymer distribution. As the equation is an integro-differential equation, the solution is non-trivial and the various solution methods, e.g. analytical methods, direct discretization, method of moments etc., are well documented in the excellent book by D. Ramkrishna (2000). In this work, only random and chain-end scission will be covered as these are the two common enzymatic scissions encountered.

2.2.1 Random Scission

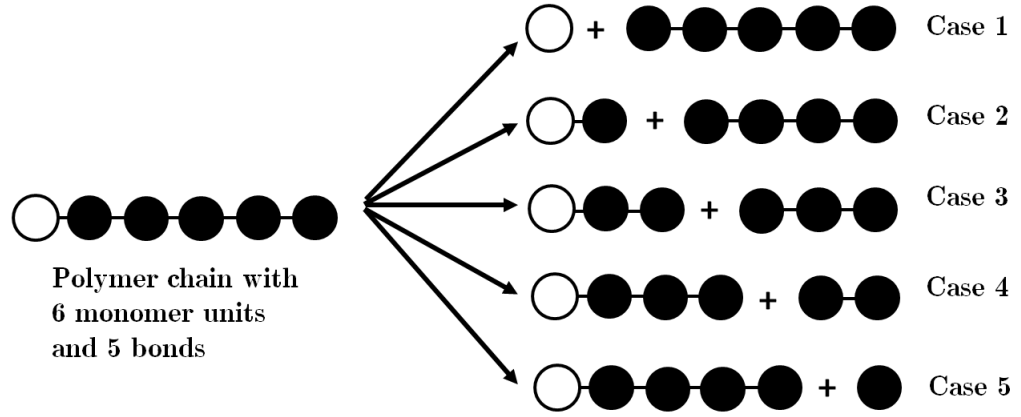


Figure 2.1: Illustration of random scission where the cases 1 to 5 are the equally probable products resulting from the scission of a polymer chain with 6 monomer units and 5 bonds.

In the case of random scission, the bonds along a particular polymer chain are hydrolyzed at random. Referring to Figure 2.1, for the scission of a polymer chain with 6 monomer units and 5 bonds, cases 1 to 5 are the equally probable products of scission. Since polymer chains are made up of discrete monomers, it is often represented by the following discrete Population Balance Equations (PBEs) (Ziff & McGrady, 1985) which are expressed in terms of the intrinsic rate of breaking an $(i + j)$ -mer into an i -mer and a j -mer, i.e. F_{ij}^α (Ziff & McGrady, 1985):

$$\frac{dc_i(t)}{dt} = 2 \sum_{j=i+1}^N c_j(t) F_{i,j-i}^\alpha - c_i(t) \sum_{k=1}^{i-1} F_{k,i-k}^\alpha \quad (2.4)$$

where $F_{i,j-i}^\alpha = k_j^\alpha b_{ij}^\alpha$ and $\sum_{k=1}^{i-1} F_{k,i-k}^\alpha = k_i^\alpha$. Here, the superscript “ α ” is used to denote random scission and c_i is the molar concentration density of polymer with DP = i (which can be used interchangeably with the molar concentration C_i since c_i equals C_i divided by a unit DP interval). The stoichiometric kernel for discrete random scission is given as:

$$b_{ij}^{\alpha} = \frac{1}{j-1} \quad (2.5)$$

This stoichiometric kernel coupled with different forms of rate kernel will result in different final expressions of Eq. (2.4). The common power law rate kernels are given as $k_i^{\alpha} = \hat{k}^{\alpha} (i-1)^n$ ($n = 0, 1, 2, \dots$) where \hat{k}^{α} is a constant. Other forms of rate kernels were also reported in the literature (Staggs, 2002, 2005, 2006). In Eq. (2.4), $i = 1, 2, 3 \dots N$, which corresponds to the DP of the monomer, dimer, trimer up to the largest DP at N . Equations (2.4) - (2.5) are the exact representation of discrete random polymer scission where scissions of polymer with all chain lengths are accounted for explicitly. Ideally, if one were to model random scission using PBM, these equations should be employed as they serve as the golden standard for discrete polymer scission. However, the order of N is usually large (Staggs, 2002, 2005, 2006), which renders the solution to such a large system of differential equations impractical. Particularly for biopolymers such as starch or cellulose, values of N spanning the order of thousands are not uncommon (Breuninger, Piyachomkwan, & Sriroth, 2009; Griggs et al., 2012a). Although analytical solutions for specific cases of discrete random binary scission had been provided by Ziff and McGrady (1985) as well as Ziff (1992), these are often restricted to certain rate kernels and initial conditions. Moreover, they are also of limited use when multiple complex depolymerization phenomena co-exist.

The impracticality of solving the discrete PBEs at realistic/large values of N calls for the development of approximate solutions at a more affordable computation cost. This is often done through the use of PBE in the continuous domain, given as the following alternative form of Eq. (2.1):

$$\frac{\partial c(v, t)}{\partial t} = 2 \int_v^{\infty} F^{\alpha}(v, w-v) c(w, t) dw - c(v, t) \int_0^v F^{\alpha}(z, v-z) dz \quad (2.6)$$

where the intrinsic rate of scission was used, i.e. $F^\alpha(v, w-v) = k^\alpha(w)b^\alpha(v, w)$ and

$\int_0^v F^\alpha(z, v-z) dz = k^\alpha(v)$. The continuous equivalent of Eq. (2.5) is:

$$b^\alpha(v, w) = \frac{1}{w} \quad (2.7)$$

For the common $k^\alpha(v)$'s, e.g. $k^\alpha(v) = \hat{k}^\alpha v^n$ ($n = 0, 1, 2, \dots$) etc., analytical solutions to Eqs. (2.6) – (2.7) (which in some cases were specific for a particular initial condition) exist and were given by Ziff and McGrady (1985). However, as alluded to above, analytical solutions are often confined to specific forms of rate kernels and initial conditions and therefore are of limited general use. Using the continuous model, the method of moments can be used to solve for the desired moments of the polymer distribution undergoing random scission (M. Kostoglou & Karabelas, 2002, 2004; D. Ramkrishna, 2000), thus bypassing the need for solving large scale discrete problems. This had been done in the past for applications in which primarily the zeroth and the first moments of the polymer distribution (indicating the total molar and mass concentrations of the polymer respectively) are variables of interest (Karmore & Madras, 2001; Madras, Chung, Smith, & McCoy, 1997; Madras, Smith, & McCoy, 1997; McCoy & Wang, 1994). Although the method of moments admits a simple and efficient solution, it is designed specifically for the extraction of important properties of the polymer distribution and not aimed at extracting the full distribution or the temporal evolution of the molar concentration of any given oligomer. The latter is especially important in the case where the concentration of oligomers is of commercial significance, thus requiring their distribution to be tracked accurately.

Another important class of solution for random scission is the use of numerical techniques by means of discretization of the continuous PBE. Other than the obvious use of the conventional finite difference, finite element or finite volume techniques, the

continuous PBE can also be solved by using the higher order methods based on approximating the full polymer distribution with sets of local/global orthogonal functions or spline polynomials (Canu, 2005; Eyre, Everson, & Campbell, 1998; Hamilton, Curtis, & Ramkrishna, 2003; Liu & Tadé, 2004; Mantzaris, 2005). The higher order methods despite being very accurate are computationally demanding (M. Kostoglou & Karabelas, 2009). For this reason, the so-called sectional method appears as an attractive alternative to solving the continuous PBE in that it is a compromise between the very accurate but computationally demanding higher order methods and the very efficient but inaccurate method of moments. Examples of sectional methods which had been employed to model random scission include the cell average technique developed by J. Kumar et al. (2008), the extended cell average technique by M. Kostoglou (2007), the method by Hill and Ng (1995), and the pivoting techniques by S. Kumar and Ramkrishna (1996a) as well as S. Kumar and Ramkrishna (1996b). Among these more recent techniques, the Fixed Pivot (FP) technique by S. Kumar and Ramkrishna (1996a) as noted in the extensive study by M. Kostoglou and Karabelas (2009) is still the state of the art. Derived from the principle of conserving any two chosen integral properties (e.g. zeroth and first moments of the distribution) through discretization of the continuous PBE, the FP technique retains the design intention of the method of moments and yet is able to predict the temporal evolution of polymer distribution to a good degree of accuracy (M. Kostoglou & Karabelas, 2009). The FP technique is also known for its ease of use and generality, which in contrast to analytical solutions can be amended for engineering purpose to accommodate simultaneously occurring complex depolymerization phenomena. As far as random scission is concerned, its solution using the aforementioned numerical techniques is rather established.

Although the continuous PBE can be used to approximate the discrete nature of polymeric scission, this is applicable only at large DPs (McCoy & Madras, 2001). Therefore, to obtain the solution to polymeric scission with reasonable accuracy over the entire DP spectrum, the solution must be formulated to cover both ends of the DP range. This can be achieved by employing a discrete-continuous approach where the low DP end is discretely formulated while the high DP end is approximated by the continuous PBE (Gelbard & Seinfeld, 1979; S. Kumar & Ramkrishna, 1996a). This strategy had been employed by S. Kumar and Ramkrishna (1996a) using the FP technique in their attempt to solve a simultaneous aggregation and breakup process with a random stoichiometric scission kernel, approximating a process with $N = 10,000$ with as few as 40 – 50 ODEs.

2.2.2 Chain-End Scission

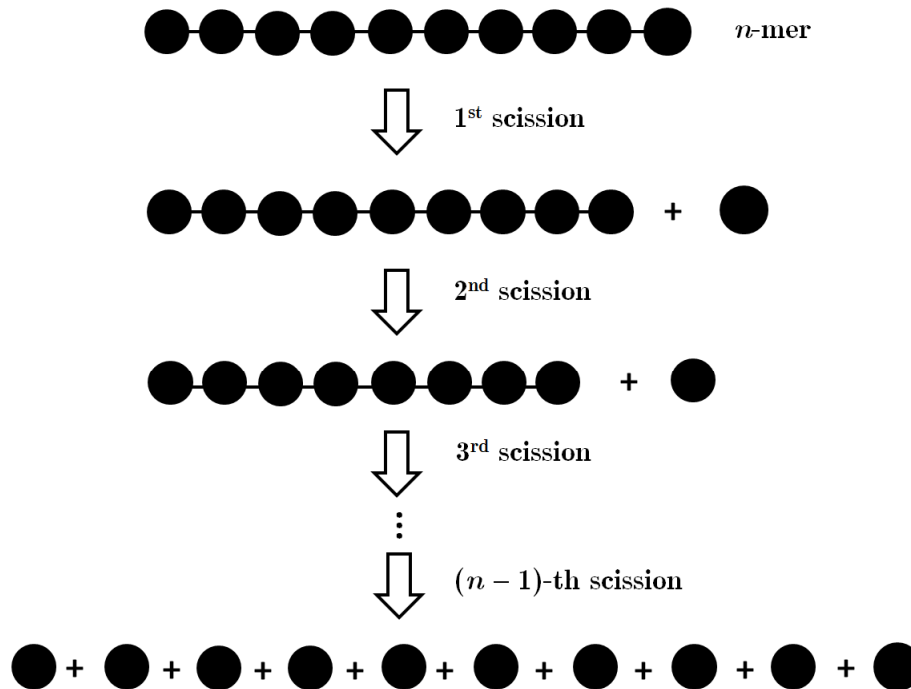


Figure 2.2: Illustration of chain-end monomer scission where monomer is removed successively from the end of the chain.

Chain-end scission is an important class of polymer degradation whereby a monomer or a dimer is removed from the end of the chain (for example, see Figure 2.2). As with the solution for random scission of polymers, the benchmark solution for chain-end scission is via the discrete form (Kostoglou, 2000). Assuming first order kinetics, the temporal evolution of the molar concentration density for every polymer species undergoing chain-end monomer scission is given as:

$$\frac{dc_1(t)}{dt} = 2k_2^\gamma c_2(t) + \sum_{j=3}^N k_j^\gamma c_j(t) \quad (2.8)$$

$$\frac{dc_i(t)}{dt} = k_{i+1}^\gamma c_{i+1}(t) - k_i^\gamma c_i(t) \quad i = 2, 3, \dots, N-1 \quad (2.9)$$

$$\frac{dc_N(t)}{dt} = -k_N^\gamma c_N(t) \quad (2.10)$$

Here, the superscript “ γ ” is used to denote chain-end scission. From Eq. (2.8), the splitting of one dimer unit to two monomer units is accounted for explicitly by the first term on the RHS. As with random scission, different rate kernels may be used to suit the specific processes described (Griggs et al., 2012a, 2012b; S. Kumar & Ramkrishna, 1996a).

To approximate the discrete model which contains large value of N , the continuous model for chain-end scission is given as:

$$\frac{\partial c(v, t)}{\partial t} = \int_v^\infty F^\gamma(v_m, w - v_m) c(w, t) dw - c(v, t) \int_0^v F^\gamma(z_m, v - z_m) dz \quad (2.11)$$

where $F^\gamma(v_m, w - v_m) = k^\gamma(w) b^\gamma(w - v_m, w)$ and $\int_0^v F^\gamma(z_m, v - z_m) dz = k^\gamma(v)$. Here, the subscript “ m ” denotes monomer, i.e. $v_m = 1$. The stoichiometric kernel for chain-end scission is:

$$b^\gamma(w - v_m, w) = \delta(v - [w - v_m]) \quad (2.12)$$

As the enforcement of Dirac delta $\delta(v - [w - v_m])$ is selective at $v = w - v_m$ and does not include the accumulation of monomer, $c(v, t)$ in Eq. (2.11) refers to the molar concentration density of the polymer population excluding monomer. The factor “2” present in Eqs. (2.1) and (2.6) is omitted for the same reason. Binary scission is thus completed by accounting for the accumulation of monomer separately:

$$\frac{\partial c_1(v, t)}{\partial t} = \int_v^\infty F^\gamma(v_m, w - v_m) c(w, t) dw \quad (2.13)$$

Here, $F^\gamma(v_m, w - v_m) = k^\gamma(w) b^\gamma(v_m, w)$, $c_1(v, t) = C_1(t) \delta(v - v_m)$ is the molar concentration density of the monomer and the corresponding stoichiometric kernel is:

$$b^\gamma(v_m, w) = \delta(v - v_m) \quad (2.14)$$

with $b^\gamma(w - v_m, w) = b^\gamma(v_m, w)$ in fulfilment of the symmetry condition. The set of Eqs. (2.11) – (2.14) is referred to as the “fundamental continuous model” for chain-end scission by McCoy and Madras (2001). The approximation of the discrete model by the fundamental continuous model is only valid if the DP of the polymer is very large (Kostoglou, 2000; McCoy & Madras, 2001). As with random scission, using the fundamental continuous model, the method of moments can be used to solve for the desired moments of the polymer distribution undergoing chain-end scission (Madras, Smith, & McCoy, 1996a, 1996b; Madras, Smith, et al., 1997; McCoy, 1999, 2001; Wang, Smith, & McCoy, 1995). Nonetheless, the method was shown recently to be appropriate only for cases where the concentration of dimer is insignificant (Stickel & Griggs, 2012). Specifically, feasible solutions using the method of moments can only be attained if the depletion of one dimer unit to form two monomer units can be neglected, otherwise it would become an ill-posed problem fraught with difficulties (Diemer & Olson, 2002; V. John, Angelov, Oncul, & Thevenin, 2007).

Due to the ease of mathematical manipulation (McCoy & Madras, 2001), another form of continuous model had also appeared in the literature for chain-end scission. Specifically, applying the Taylor series expansion on the fundamental continuous model resulted in a first order hyperbolic PDE. Truncating after the first order term, Kostoglou (2000) gave an analytical treatment of this approximate continuous model and showed that it is only valid on the condition that the number-average DP is much larger than one. In addition, a discrete-continuous mixed approach (Gelbard & Seinfeld, 1979; S. Kumar & Ramkrishna, 1996a) was also proposed where the analytical solution for the larger DPs was coupled to a discrete set of equations relating the smaller DPs. The use of discrete equations for the smaller DPs thus eliminates the dimer issue alluded to above, as this can now be explicitly accounted for. The analytical study by Kostoglou (2000) gave valuable insights on chain-end scission but the results are not easily extensible to include more complex depolymerization phenomena, e.g. a combination of random and chain-end scission (Griggs et al., 2012a, 2012b). Recently, Stickel and Griggs (2012) extended the approximate continuous model by including the second order derivative term of the Taylor series expansion and solved the resulting PDE by finite difference. Although good agreement to the discrete model was obtained in the polymer distribution, no consideration was made to ensure the conservation of moments, which could be the primary concern in some applications (S. Kumar & Ramkrishna, 1996a; Madras, Chung, et al., 1997; Oyerokun & Vaia, 2012; Rangarajan, Bhattacharyya, & Grulke, 1998; Striegel, 2003).

In contrast to random scission where its solution using the sectional techniques had been thoroughly explored, e.g. in the studies by Hill and Ng (1995), S. Kumar and Ramkrishna (1996a), S. Kumar and Ramkrishna (1996b), J. Kumar et al. (2008) etc., this is not so for chain-end scission. Particularly for the popular FP technique, only a glimpse

of its performance on chain-end scission could be found in the work of Vanni (2000). In that work for modeling aggregation-breakage processes, out of the many stoichiometric kernels tested, chain-end scission was included as a limiting case to discriminate the performance of the sectional methods (including the FP technique). However, that work paid no attention to the dimer issue alluded to above or to extracting the temporal evolution of individual oligomer concentration. Detailed guidelines e.g. on meshing or the range of applicability, were also absent. As the potential of the FP technique to handle chain-end scission appears not to have been thoroughly explored, a part of this work will be dedicated to formulate the solution to chain-end scission using the FP framework. Success in doing so will enable the simulation of multiple co-existing complex depolymerization phenomena using a single solution framework.

2.2.3 Population Balance Modelling of Saccharification Processes

Insofar as the use of PBM have yet to be found in the SSF literature, its employment for pure enzymatic scission of natural polymers had been sparingly reported. In the work of Chang et al. (2002), a continuous population balance model was developed for the hydrolysis of wheat starch by α -amylase. Random scission was assumed with a first order rate kernel [$k^\alpha(T, t) = \hat{k}^\alpha(T) E_\alpha(t) C_g(t)$; T = temperature, E_α = concentration of enzyme exhibiting random scission (α -amylase), C_g = concentration of gelatinized starch] and the PBE was solved using the method developed by Hill and Ng (1995) using a geometric discretization covering the approximate range of DP from 2 – 9.3×10^4 with 110 ODEs. The study was concerned mainly with predicting the performance of a viscosity-based device, and thus the validity of the prediction with regards to the concentration density as well as the oligomer concentration was not pursued. In another study, a discrete population balance model for the action of endo-glucanase in randomly

hydrolyzing cellulose was proposed (Hosseini & Shah, 2011b). The discrete rate kernel employed was $k_i^\alpha = \hat{k}^\alpha v_i^2$ where v_i is the DP of polymer chains with i number of monomers. The maximum DP for the cellulose population was chosen as $N = 1200$. As N here is not very large, the fully discrete PBEs were solved directly. This work was extended to include the action of exo-glucanase in their subsequent work (Hosseini & Shah, 2011a). Equal reactivity was assumed for the successive removal of cellobiose (maximum DP = 4000) from the non-reducing ends of cellulose chains, i.e. $k_i^\gamma = \hat{k}^\gamma$. Following this, a model which includes the combined action of endo- and exo-glucanases was also proposed. Nevertheless, the resulting model equations were not solved by the authors. A continuous model for the combined action of the endo- and exo-glucanases was proposed in the work of Griggs et al. (2012a) and Griggs et al. (2012b), each focusing on different aspects of the hydrolysis biochemistry. In their studies, equal reactivity was assumed for the chain-end scission by exo-glucanase. However, an enzyme complexation step was also included in their proposed mechanism and the rate kernel used was $k^\gamma(v) = \hat{k}^\gamma/v$ due to the hypothesis that the time required to find a reducing end is proportional to the chain length. Their formulation for random scission by endo-glucanase employed a normalized rate kernel of $k^\alpha(v) = \hat{k}^\alpha (v/\overline{M}_n)$ where \overline{M}_n is the number-average DP. Central finite difference method was used to solve the resulting system of PBEs and the equations involving chain-end scission were formulated using the method of Stickel and Griggs (2012). Most of the studies above were mainly concerned with predicting the distribution of the polymer populations and the concentrations of the oligomers analyzable by current chromatographic techniques were not explicitly displayed. In the context of SSF, as microbes are known to metabolize the oligomers, further clarification in the PBM framework is needed to establish a linkage between the relevant oligomer with the metabolism of microbes.

2.3 Cybernetic Modelling (CM)

The cybernetic perspective of microbial growth was first introduced by D. Ramkrishna (1983). The framework establishes its premises upon the fact that microbial systems when facing diverse environmental changes are capable of regulating their growth processes in order to optimize growth. The existence of such regulatory mechanisms was reported as early as the year 1942 in the classic study by Monod (1942) where the preferential consumption of one sugar over the other by the bacterial culture resulted in a diauxic growth pattern. Employing the traditional kinetic framework for modelling biological systems, such a complex growth pattern can only be predicted if the substrate preference is supplied *a priori* as an input to the model. Moreover, the incorporation of metabolic regulation into the traditional kinetic framework increases the model complexity enormously and is often impeded by the scarcity of mechanistic details (D. Ramkrishna & Song, 2012). In the face of these challenges, the CM framework has taken another route in which metabolic regulation is attached to a goal-seeking behaviour of the organism. With this in mind, the CM framework adopts a simpler mathematical description of the biological systems which are capable of metabolic regulation by associating their response to a certain control policy.

2.3.1 Cybernetic Modelling Equations

Since its first inception, the CM framework has evolved tremendously over the past three decades, progressing from the simple microbial growth models built upon gross biochemical pathways to its current capability in utilizing larger scale metabolic network for metabolic engineering purposes. A comprehensive review of the development of the CM framework can be found in the paper by D. Ramkrishna and Song (2012). A

simplified representation of the CM machinery is shown in Figure 2.3. The metabolic enzymes (e_i) play a critical role in the cellular metabolism of different substrates (\mathcal{N}_i) present in the abiotic phase. Once the substrates enter the biotic phase, they go through a series of biochemical conversion with the release of intracellular metabolites (ϕ), in which some are ultimately excreted into the external environment as products (\mathcal{P}). Meanwhile, cells grow and the biomass (X) increases. The CM framework views the levels and activities of the metabolic enzymes as being regulated systemically through the cybernetic variables U_i and V_i . Conceptually, one could view U_i as the fractional allocation of cellular resources to produce e_i , and V_i as the extent e_i is activated. The CM equations, while present in numerous variants, can be expressed in the following general system of nonlinear ODEs:

$$\frac{1}{X} \frac{d\mathcal{N}}{dt} = \mathbf{W}_{\mathcal{N}} \mathbf{r}(V, \mathbf{e}, \mathcal{N}) \quad (2.15)$$

$$\frac{1}{X} \frac{d\mathcal{P}}{dt} = \mathbf{W}_{\mathcal{P}} \mathbf{r}(V, \mathbf{e}, \mathcal{N}) \quad (2.16)$$

$$\frac{d\phi}{dt} = \mathbf{W}_{\phi} \mathbf{r}(V, \mathbf{e}, \mathcal{N}) - \mu \phi \quad (2.17)$$

$$\frac{1}{X} \frac{dX}{dt} = \mu \quad (2.18)$$

$$\frac{d\mathbf{e}}{dt} = \mathbf{p}_e + \mathbf{D}(U) \mathbf{r}_e - [\mathbf{D}(\boldsymbol{\beta}) + \mu \mathbf{I}] \mathbf{e} \quad (2.19)$$

where X is the mass concentration of biomass, μ is the total specific growth rate, and \mathcal{N} is the concentrations of extracellular substrates, \mathcal{P} is the concentration of extracellular products, and ϕ is the concentration of intracellular metabolites. The matrices $\mathbf{W}_{\mathcal{N}}$, $\mathbf{W}_{\mathcal{P}}$ and \mathbf{W}_{ϕ} contain the stoichiometry of the reactions. The symbol \mathbf{r} is the regulated fluxes (or rates defined per unit of biomass), \mathbf{p}_e is the constitutive enzyme synthesis rates, \mathbf{r}_e is the inducible enzyme synthesis rates, $\boldsymbol{\beta}$ is the enzyme degradation rate constants and \mathbf{e}

is the enzyme levels. Unless dimensionless, units of concentration related terms may be expressed either on a molar or a mass basis wherever appropriate. The matrix \mathbf{I} is the identity matrix and the operator $\mathbf{D}(\bullet)$ converts the input vector into a diagonal matrix such that:

$$\mathbf{D}(\boldsymbol{\beta}) = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \beta_{n_e} \end{bmatrix}; \quad \mathbf{D}(\mathbf{U}) = \begin{bmatrix} U_1 & 0 & \cdots & 0 \\ 0 & U_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_{n_e} \end{bmatrix}; \quad (2.20)$$

$$\mathbf{D}(\mathbf{V}) = \begin{bmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & V_{n_e} \end{bmatrix}$$

where n_e is the size of \mathbf{e} .

The general form of \mathbf{U} and \mathbf{V} can be written as (Song et al., 2009):

$$\mathbf{U} = \frac{\mathbf{R}^+}{\sum_k \mathbf{R}_k^+} \quad (2.21)$$

$$\mathbf{V} = \frac{\mathbf{R}^+}{\max(\mathbf{R}^+)} \quad (2.22)$$

where $\mathbf{R}^+ = [\mathbf{R}_1^+ \quad \mathbf{R}_2^+ \quad \cdots \quad \mathbf{R}_k^+]^T$ with $\mathbf{R}_k^+ = \max(\mathbf{R}_k, 0)$ and \mathbf{R}_k is the return on investment from the k -th alternative. The form of \mathbf{R}_k is dependent on the formulation of the metabolic objective function (Song et al., 2009). As the ability to reproduce is important for ensuring survival, the biomass growth rate had been frequently employed as the objective function (Kompala & Ramkrishna, 1986; Kompala, Ramkrishna, & Tsao, 1984). Other choices of objective function had also been employed, e.g. the substrate uptake rates (Turner, Ramkrishna, & Jansen, 1989). Defining the metabolic objective function over a finite time interval $(t, t + \Delta t)$ as $\Delta J = J(t + \Delta t) - J(t)$, the cells ensure their survival by manipulating the cybernetic variables according to the following maximization problem:

$$\begin{aligned}
& \max_{U, V} \Delta J \\
& \text{s.t. } \|U\|_1 = 1, U_k \geq 0, V_k \leq 1
\end{aligned} \tag{2.23}$$

By setting $\Delta t = 0$, the control policy collapse to the classic forms determined by the ‘Matching and Proportional laws’ (Young & Ramkrishna, 2007). The constraint of $\|U\|_1 = 1$ is due to the fact that U is associated with the fractional allocation of a common critical resource for the synthesis of enzymes for different cellular reactions. The elements in V need not sum to 1 as they are not viewed based on the sharing of a common resource. Although the CM equations given above can be used to predict the effect of multiple substrates on microbial growth, the existing framework does not cater explicitly to systems where the excretion of the extracellular depolymerases is required to break large polymers into consumable substrates.

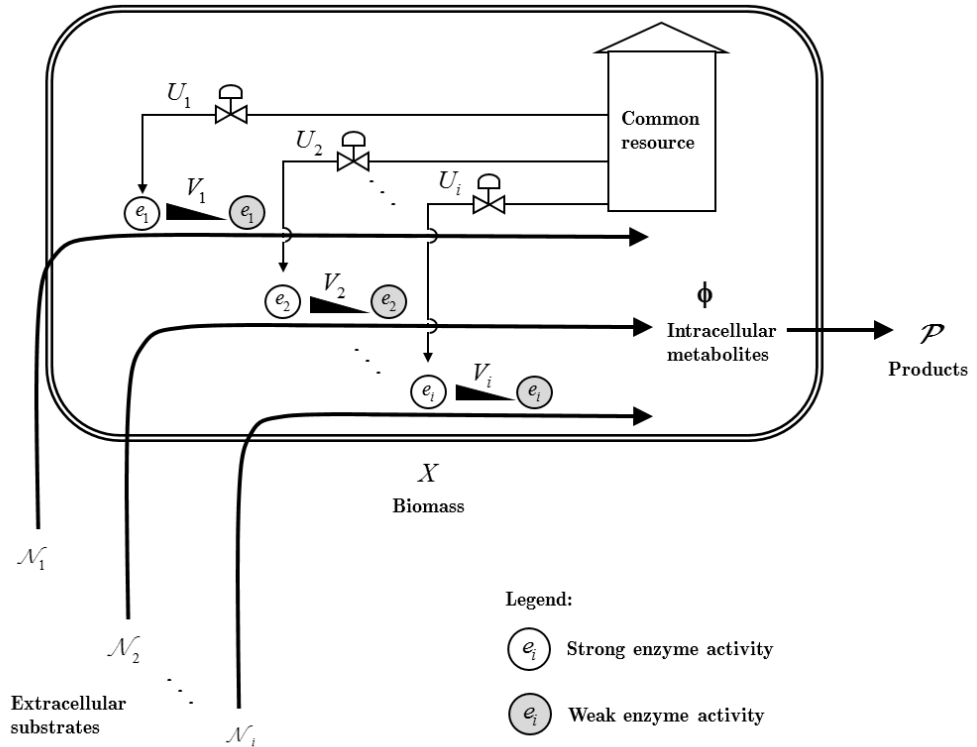


Figure 2.3: Simplified diagram illustrating the different components of the CM framework. The biotic phase is represented by the region enclosed within the dashed line.

2.3.2 Cybernetic Modelling for SSF Processes

The use of the CM framework to model cellular response to complex nutrient environment had been very successful. At its early beginnings, the CM framework developed by Kompala and Ramkrishna (1986) was used to predict the growth pattern of *Klebsiella oxytoca* in batch cultures growing on a mixed substrates formed from combinations of glucose, xylose, arabinose, lactose and fructose. The more advanced developments, e.g. that which is found in the study by Geng et al. (2012), saw the enhanced capability of the framework to predict the co-fermentation of a multiple sugar mixture by multiple species. The literature on CM had been mainly focused on assimilation of sugars which can be directly consumed, e.g. glucose, xylose, mannose etc. However, in the case of SSF, the starting raw material is usually natural polymers which cannot be directly consumed unless further degraded to smaller products. Thus, a mechanism for regulating the levels and activities of the extracellular enzymes necessary for degrading the larger substrates must be in place for proper predictions to be made. This aspect of SSF was first embedded in the CM framework in the work of Gadgil et al. (1996), where the yeast employed was unable to consume melibiose (a di-saccharide) directly but that extracellular α -galactosidase must be excreted to hydrolyze melibiose into glucose and galactose. The study further assumed that glucose represses the production of α -galactosidase only if its concentration exceeds a critical threshold. In addition, the concentration of α -galactosidase was assumed to correspond to the key enzyme in galactose metabolism. Taking a cue from this study, Altintas et al. (2002) employed a similar strategy to model the SSF of starch by a recombinant yeast which excretes a fusion protein which exhibits both α -amylase and glucoamylase activity. In this case, starch was assumed to be hydrolyzable to form reducing sugars and glucose, where the concentration of the extracellular hydrolytic enzyme corresponded to the key

enzyme in the metabolism of reducing sugars. Their model predicted a sluggish response for starch degradation even after the initial glucose loading had been consumed, which implied an unsatisfactory representation of the enzyme induction/repression process. Following this, a more expanded pathway for starch degradation and assimilation of the corresponding hydrolysates was proposed by Ochoa et al. (2007). Their model was constructed based on the cybernetic principles elaborated by Varner and Ramkrishna (1999b) and the performance of the model was relatively better than that proposed by Altintas et al. (2002). Nonetheless, both the models by Altintas et al. (2002) and Ochoa et al. (2007) were not validated for the case where starch is present as the sole initial substrate. Other attempts employed the CM framework to model SSF processes in which the enzymes required for saccharification were externally dosed and not produced *in situ* by the microbes (Chavan et al., 2009; Ko et al., 2010; Ganti S. Murthy et al., 2012). To this end, all the models above were more elaborate in describing the kinetics of fermentation while attempting to represent the kinetics of hydrolysis with crude M-M type expressions. This study attempts to address this inadequacy by proposing to integrate the PBM and the CM principles in a single framework, thus enabling a more fundamental analysis of the SSF phenomena.

CHAPTER 3 : RESEARCH METHODOLOGY

3.1 Preamble

As the general nature of this work involved detailed mathematical derivations, for ease of reading, such materials were placed in their corresponding chapters preceding the presentation of the results and discussions. Therefore, this chapter gives only the general methodology and line of thought involved in executing the various parts of this work. The detailed formulations are found in the respective chapters. All simulations involved in this work were programmed using MATLAB[®] program R2009b. When dealing with a system of ODEs, the ‘ode45’ sub-routine was always the “first preference” as per the recommendation from MATLAB (Houcque, 2008). Should the ‘ode45’ become inefficient, the sub-routine ‘ode15s’ was used instead. Where appropriate, such information will be mentioned.

In addition, three different workstations were used throughout the course of this study, namely workstations A, B, and C respectively. Workstation A comprises of an Intel[®] Core[™]2 Duo CPU T8100 with a clock speed of 2.10 GHz and 3GB of installed memory (RAM), whereas workstation B comprises of an Intel[®] Xeon[®] Processor W3520 with a clock speed of 2.66 GHz and 8GB of RAM. Finally, workstation C comprises of Intel[®] Core[™] i7-4700HQ2 CPU with a clock speed of 2.40 GHz and 16GB of RAM. Where the specifications of the workstation become important such as when the absolute computational time is reported, the particular workstation used will be mentioned.

3.2 General Workflow

The general workflow of the research is summarized in Figure 3.1 below. Details of the methodologies employed are presented in the subsequent sections.

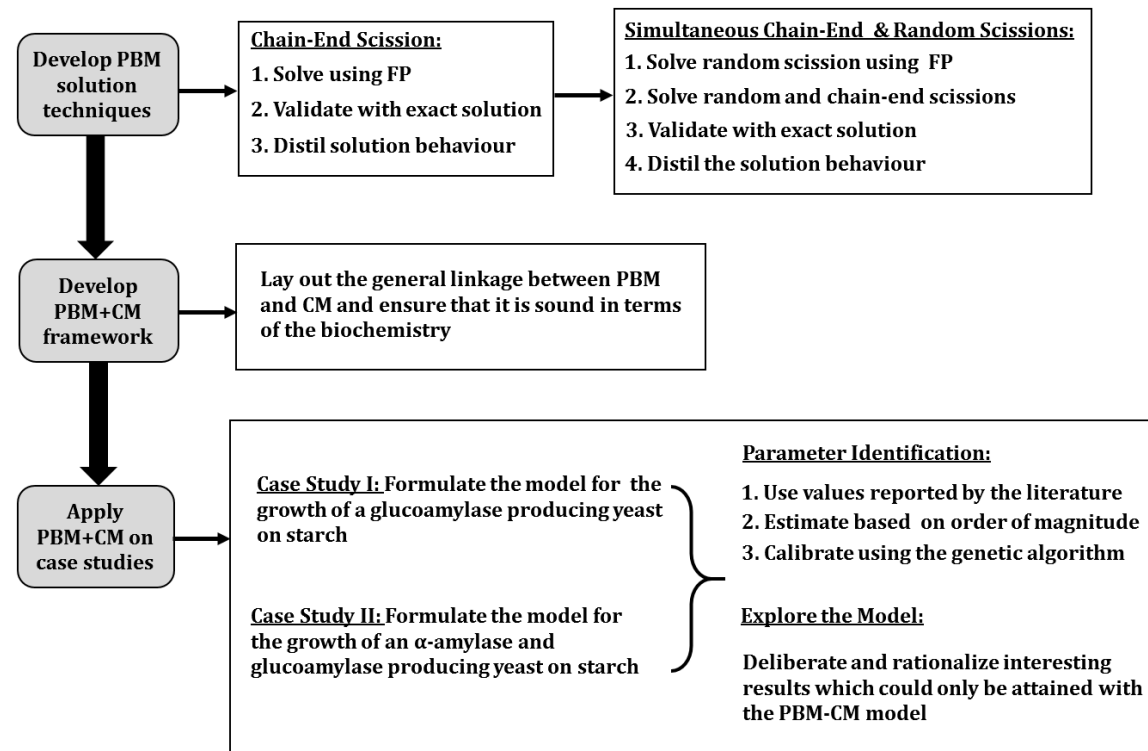


Figure 3.1: Flowchart of the research workflow

3.3 Development of the Numerical Technique for Solving Population Balance Equations

Two techniques were developed as far as the solution technique for solving PBEs is concerned. The first is the development of the FP technique in solving chain-end scission, which as alluded to in Section 2.2.2, had not been fully explored previously. Second, the FP solution for random scission was combined with the FP solution for chain-end scission to enable the solution of simultaneously occurring random and chain-end scissions. Detailed formulations regarding the solution of the PBEs using the FP technique are given in Chapter 4. Nonetheless, the key methodologies employed are given below.

3.3.1 Identification of the Relevant Stoichiometric Kernels, Rate Kernels and Initial Distribution for Solving the PBEs

Regardless of the technique used to solve PBEs, three critical items that are always required for solving PBEs are: a) the stoichiometric kernel, b) the rate kernel, and c) the initial polymer distribution. Selection of the stoichiometric kernel was done by considering two points. First, a survey of the literature concerning the common form of enzymatic scissions was done. For the hydrolysis of natural polymers, e.g. starch, cellulose etc., two major classes of enzymatic actions were uncovered to be prominent in the breakdown of these natural polymers. Briefly, the depolymerases could be generally divided into those that are endo-acting as well as those which are exo-acting (Bansal et al., 2009; Robyt, 2009). Second, a search in the pertinent population balance literature on enzymatic hydrolysis of natural polymers revealed that the action of endo-acting as well as the exo-acting enzymes can be appropriately modelled with the random scission kernel

and the chain-end scission kernel respectively (Chang et al., 2002; Griggs et al., 2012a, 2012b; Hosseini & Shah, 2011a, 2011b).

Upon selecting the appropriate stoichiometric kernel, the rate kernel and the initial distribution had to be determined in order to solve the PBEs. In the course of developing the solution technique, although different forms of rate kernels had been reported in the literature (cf. Section 2.2.3) for the hydrolysis of natural polymers, the classic power law rate kernels (S. Kumar & Ramkrishna, 1996a) were selected for simplicity as the objective here is to mainly develop the solution technique. As for the initial distribution of the polymers, two different forms of expressions were used throughout this work. The first is a Schulz-Zimm distribution (Dole, 1972), which had been used in the work of Chang et al. (2002) to model starch. This distribution is empirical in nature and is commonly used for describing condensation polymers with a skewed long tail at the high molecular weight end. The Schulz-Zimm distribution is given as:

$$c(v, 0) = \frac{m_s(0)h(v)}{\int_1^N [(162v+18)-180]h(v)dv} \quad (3.1)$$

where $m_s(0)$ is the initial mass concentration (mass/volume) of the polymer, v is the continuous DP, N is the maximum DP of the polymer, and the remaining quantities are given by:

$$h(v) = \left(\frac{v-1}{\kappa}\right)^{\omega-1} \exp\left(-\frac{v-1}{\kappa}\right) / [\kappa\Gamma(\omega)] \quad (3.2)$$

$$\omega = \frac{\overline{M}_n}{\overline{M}_w - \overline{M}_n}, \quad \kappa = \overline{M}_w - \overline{M}_n, \quad \Gamma(\omega) = \int_0^\infty t^{\omega-1} \exp(-t) dt \quad (3.3)$$

In Eqs. (3.3), \overline{M}_n is the number-average DP, \overline{M}_w is the weight-average DP and $\Gamma(\omega)$ is the gamma function. Other than the Schulz-Zimm distribution, the other form of initial distribution used in this work has the following form:

$$c(v, 0) = \frac{c_{in}}{\kappa^\omega \Gamma(\omega)} v^{\omega-1} \exp\left[-\frac{v}{\kappa}\right] \quad (3.4)$$

where c_{in} is a constant. This form of initial distribution is also common in the literature (Griggs et al., 2012a, 2012b; Stickel & Griggs, 2012). The distribution function shown in Eq. (3.4) is similar to that shown in Eq. (3.2) except that the values of ω and κ in Eq. (3.2) were correlated with \overline{M}_n and \overline{M}_w .

3.3.2 Code Validation

Be it the programming of the fully discrete (exact) or the FP solution, it is important to ensure that the code correctly transcribes the system of equations. For the PBEs describing pure scission, the mass of the polymers should always be conserved. Therefore, to check that the code has been programmed correctly, the temporal evolution of the first moment of the polymer distribution can be used. Given that N is the maximum DP of the polymer distribution and $p+q$ is the number of discretized grid points (also called pivots), the first moment of the distribution can be calculated in the following manner:

Exact solution:

$$C^{(1)}(t) = \sum_{i=1}^N y_i C_i(t) \quad (3.5)$$

FP solution:

$$C^{(1)}(t) = \sum_{i=1}^{p+q} x_i C_i(t) \quad (3.6)$$

In Eqs. (3.5) and (3.6), $C^{(1)}$ represents the first moment of a polymer distribution, y_i is the DP of the i -th pivot for the exact solution, and x_i is the DP of the i -th pivot for the FP solution. As will be shown later in Chapter 4, $p+q$ is the total number of pivots for the FP solution, which is much smaller than N . If the PBEs had been coded correctly, $C^{(1)}(t)$ is

a constant. Upon validating that the code has been programmed correctly, the next step is to validate the FP solution with the exact solution, as discussed next.

3.3.3 Validation with the Exact (or Fully Discrete) Solution

For benchmarking purposes, the resulting temporal evolutions of the molar concentration density, the molar concentrations of selected oligomers as well as selected properties of the distribution obtained using the FP technique were compared against those of the exact solution. A few of the following quantities were used throughout. First, the dimensionless time (θ) is given by:

$$\theta = \frac{t}{t_{99\%}} \quad (3.7)$$

where $t_{99\%}$ is the time required for the molar concentration of the monomer to reach 99% of its final value. When comparing the FP solution with the exact solution, it is always important to make comparison at the late phase of the reaction, i.e. at $\theta \geq 1$ as the error accrued over time is most significant during this phase. If the FP solution is accurate even at $\theta \geq 1$, one can then conclude that the numerical approximation is appropriate for use even at the late phase of reaction where essentially all polymers had been hydrolyzed to monomers.

Second, to quantify deviations from the exact solution, the following global error indicator was used throughout this study:

$$\varepsilon_g = \frac{1}{H} \sum_{j=1}^T \frac{|C^{Exact}(t_j) - C^{FP}(t_j)|}{1 + C^{Exact}(t_j)} \quad (3.8)$$

where H is the total number of time steps. The benefit of ε_g is that it will reflect the average relative error or the average absolute error depending on the magnitude of

$C^{Exact}(t_j)$. It does not blow up artificially even with very small values of $C^{Exact}(t_j)$ often encountered here.

Finally, the error in the initial mass of polymers due to discretization was calculated as:

$$\varepsilon_S = |\hat{m}_S(0) - m_S(0)|; \quad \hat{m}_S(0) = \sum_{i=1}^{p+q} M_i C_i; \quad M_i = 162x_i + 18 \quad (3.9)$$

In Eq. (3.9), the form of molecular weight (M_i) given is applicable to glucose polymers.

3.4 Development of the Interlinked PBM and CM Framework

The core of this work is to give the general framework for interlinking the PBM and CM followed by two case studies to showcase the attractive features of such an approach in modelling SSF systems. The first case study involves the growth of a glucoamylase producing recombinant yeast on starch. This case study was chosen to mainly introduce the general capability of the framework. The second case study involves the growth of an α -amylase and glucoamylase producing recombinant yeast and was chosen to show how the framework could be used to abstract the very complicated scenario of two depolymerases attacking a common polymeric substrate. Discretization of the PBEs involved was done through the FP technique formulated in Chapter 4 of this thesis and the detailed formulations of the interlinked PBM-CM framework are given in Chapter 5. The common methodologies employed are given below.

3.4.1 Identification of Relevant Stoichiometric Kernels, Rate Kernels and Initial Distribution for Solving PBEs

In line with the discussions in Section 3.3.1, the stoichiometric kernels to be employed are none other than those for the random and the chain-end scissions. Therefore, the case study examples chosen must appropriately reflect the use of these two stoichiometric kernels. As such, referring to Figure 3.1, two case studies were selected, one of which involves the growth of a recombinant yeast capable of excreting glucoamylase (chain-end scission), while the other is capable of excreting both α -amylase (random scission) and glucoamylase (chain-end scission).

For the rate kernels, since the biochemistry of enzymatic hydrolysis cannot be neglected in this case, the use of the classic power law kernels may not be appropriate. As starch was selected as the natural polymer in the case studies, a literature search was performed to uncover the possible relations between the length of the polymer chain and its corresponding rate of hydrolysis. It turns out that the subsite theory for glucoamylase and α -amylase (Hiromi, 1970; Hiromi, Nitta, Numata, & Ono, 1973; Iwasa, Aoshima, Hiromi, & Hatano, 1974) is relevant and was thus employed. For the initial distribution of starch, the Schulz-Zimm distribution, i.e. Eqs. (3.1) - (3.3), was used primarily because it could be used to relate to starch properties such as \overline{M}_n and \overline{M}_w .

3.4.2 Identification of Model Parameters

The models developed for the case studies were calibrated based on the data reported in the literature. Specifically, the data provided by Nakamura et al. (1997) and Ülgen, Saygılı, Önsan, and Kırdar (2002) were used for case studies I and II respectively.

Extraction of the experimental data was done using the free online application, i.e. WebPlotDigitizer developed by Rohatgi (2012). Identification of model parameters was generally done through order of magnitude estimation, inference from reported values or nonlinear parameter estimation. All nonlinear parameter estimations were done using the Genetic Algorithm (GA) ('ga' subroutine) from the Genetic Algorithm and Direct Search Toolbox™ of MATLAB® R2009b. The population size of the algorithm was chosen as five times the number of parameters (Cox, 2005) as increasing it further yielded similar results at the expense of much longer computational time. The maximum number of generation was set as 200 and the maximum number of stall generations was set as 100. The remaining settings were kept as default.

As alluded to above, the determination of the values of some of the model parameters requires the use of nonlinear parameter estimation. For those not requiring nonlinear parameter estimation, the rationale for their case by case selection will be given in the relevant section in Chapter 5. For those requiring nonlinear parameter estimation, all the parameters to be estimated were normalized in the following way:

$$\Theta_i = \frac{\Theta_i - \Theta_{i,\min}}{\Theta_{i,\max} - \Theta_{i,\min}} \quad (3.10)$$

where Θ_i is the i -th parameter to be calibrated and $[\Theta_{i,\min}, \Theta_{i,\max}]$ is the range of lower and upper bounds for Θ_i . The nonlinear parameter estimation problem was formulated as:

$$\begin{aligned} \min_{\Theta} J_{opt} \\ \text{Subject to } 0 \leq \Theta \leq 1 \end{aligned} \quad (3.11)$$

where J_{opt} is the objective function and is given as:

$$J_{opt} = \sum_i \bar{W}_i \left\{ \frac{1}{\tau_i} \left[\frac{\hat{y}_i - \tilde{y}_i}{\max(\tilde{y}_i)} \right]^T \left[\frac{\hat{y}_i - \tilde{y}_i}{\max(\tilde{y}_i)} \right] \right\} \quad (3.12)$$

Here, i refers to the variables involved in calibration (e.g. biomass, starch etc.), $\hat{\mathbf{y}}_i$ is the vector of predicted output, $\tilde{\mathbf{y}}_i$ is the vector of experimental observations, τ_i is the number of observations for the variable i and \bar{W}_i (usually having values between 1 – 10) is the weight assigned to the different variables. When $\bar{W}_i = 1$, the objective function weighs all quantities of different scales and different number of points equally. In calculating J_{opt} , if the elements in $\tilde{\mathbf{y}}_i$ do not fall on the same time points as $\hat{\mathbf{y}}_i$, the ‘interp1’ sub-routine in MATLAB® R2009b was used to interpolate $\hat{\mathbf{y}}_i$.

The general workflow for the nonlinear parameter estimation using the GA is given in Figure 3.2. In assigning values to model parameters, it is important that the final values of the parameters make biological sense. Therefore, appropriate bounds for the model parameters, i.e. $[\Theta_{i,min}, \Theta_{i,max}]$, should be established. This was done by consulting the literature on past similar models which had been developed. As an example, the typical value of $\mu_{max,l}$ (maximum specific growth rate on glucose) for yeast lies in the range of [0.1, 1] (Altintas et al., 2002; Gadgil et al., 1996; Jang & Chou, 2013; Kobayashi & Nakamura, 2003, 2004). Thus, it would be impractical to set the bounds on $\mu_{max,l}$ beyond this range. Upon establishing the bounds for the model parameters, the GA was run without a known set of initial population and using unity weights in the objective function. As the GA is stochastic in nature, this might return different results on every different run, and the optimized parameters should be examined manually to ensure that they make biological sense. Once a particular set of optimized parameters was deemed ‘biologically appropriate’, this set of parameter values was used as the initial population for the next calibration procedure using the GA while still maintaining the unity weights on the objective function. Apart from examining the value of the objective function, the resulting goodness of fit was evaluated by manually scrutinizing the fitted curves as well as by

ensuring that the simulated profiles of other unfitted variables are to the best of knowledge biological sound. If some fits are relatively poorer than the others, the optimization by GA is invoked again by manually adjusting \bar{W}_i . Generally, quantities with poorer fits require larger values of \bar{W}_i than the better ones to increase the sensitivity of the optimizer toward the particular variable. The procedure is terminated when good fitting is obtained.

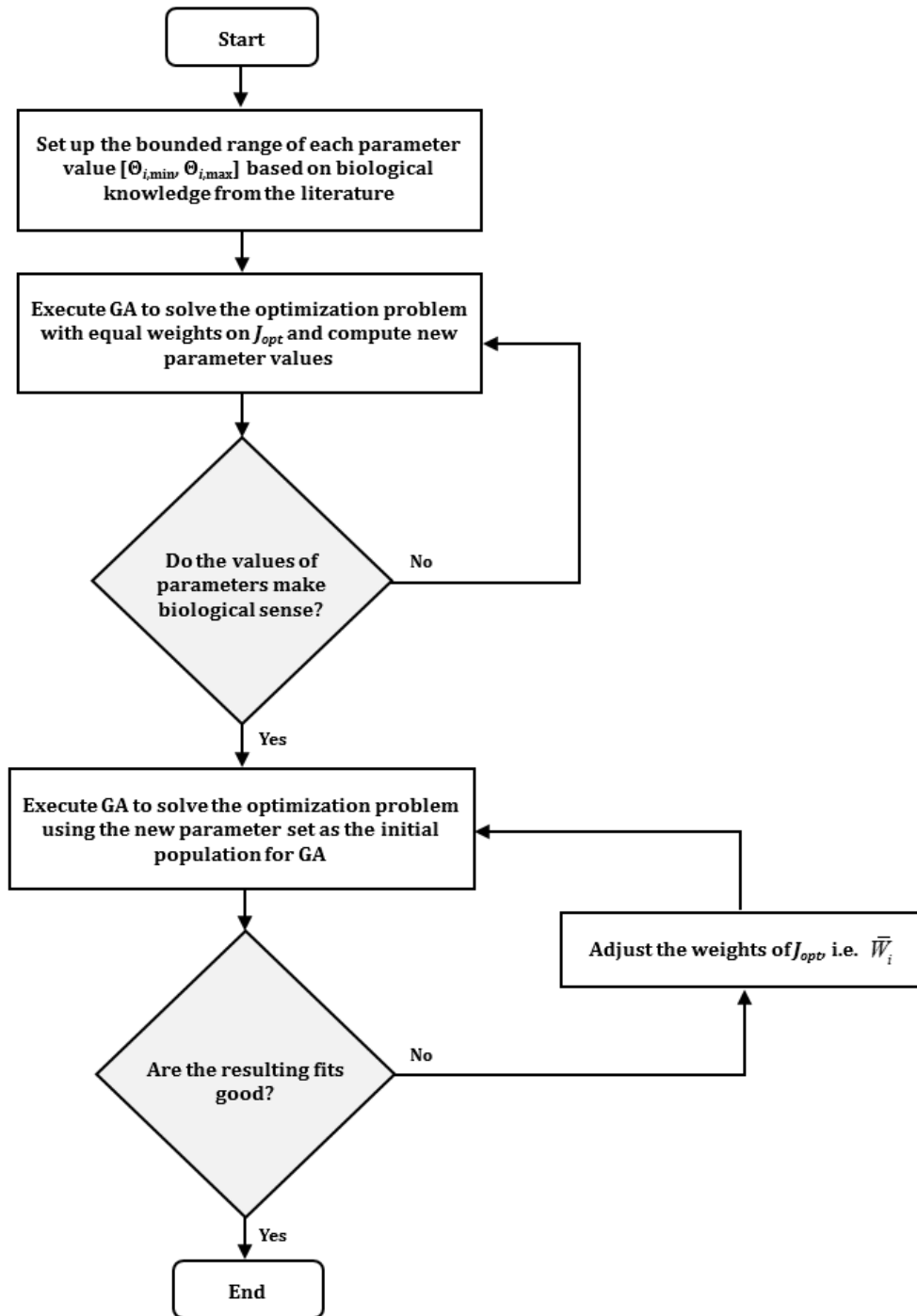


Figure 3.2: The general workflow of nonlinear parameter estimation using the Genetic Algorithm (GA).

CHAPTER 4 : MODELLING POLYMERIC SCISSIONS USING THE FIXED PIVOT TECHNIQUE

4.1 Modelling Chain-End Scission

As alluded to in Section 2.2.2, the FP technique appears to not have been thoroughly explored for solving chain-end scission. This is the aim of Section 4.1 where the FP technique was modified to account for chain-end scission in a systematic manner, thus paving the way for the ultimate description of simultaneously occurring random and chain-end scissions as the ultimate goal of this thesis.

4.1.1 Fully Discrete (Exact) Solution for Chain-End Scission

The fully discrete solution for chain-end scission had been given previously in Section 2.2.2 by Eqs. (2.8) – (2.10). Nevertheless, as the fully discrete solution is the benchmark for validating the accuracy of the FP solution which will hereafter be formulated in the following sections, the equations are reproduced here for ease of reference. For a polymer with a maximum DP = N , the discrete solution to chain-end scission consists of N ODEs. Assuming first order kinetics, the accumulation of the molar concentration of each species with time is given as:

$$\frac{dC_1(t)}{dt} = 2k_2^\gamma C_2(t) + \sum_{j=3}^N k_j^\gamma C_j(t) \quad (4.1)$$

$$\frac{dC_i(t)}{dt} = k_{i+1}^\gamma C_{i+1}(t) - k_i^\gamma C_i(t) \quad i = 2, 3, \dots, N-1 \quad (4.2)$$

$$\frac{dC_N(t)}{dt} = -k_N^\gamma C_N(t) \quad (4.3)$$

Although the analytical solutions to the above were given by Kostoglou (2000), they cannot be readily extended to cases where additional terms exist on the RHS of the ODEs, e.g. to account for other concurrent processes such as random scission.

4.1.2 Fixed Pivot Discretization for Chain-End Scission

The continuous PBE for chain-end scission presented in Eqs. (2.11) - (2.12) of Section 2.2.2 can be rewritten in the following form:

$$\frac{\partial c(v, t)}{\partial t} = \int_v^\infty k^\gamma(w) \delta(v - [w - v_m]) c(w, t) dw - k^\gamma(v) c(v, t) \quad (4.4)$$

Discretizing along the v (or continuous DP) dimension, the FP technique assumes that the population of polymers in the i -th interval is concentrated only at a chosen representative DP (called the pivot, x_i), i.e. $v_i < x_i < v_{i+1}$. The molar concentration density which excludes the monomer is thus mathematically expressed by:

$$c(v, t) = \sum_{i=2}^{\infty} C_i(t) \delta(v - x_i) \quad (4.5)$$

Here, C_i is the molar concentration for the i -th interval:

$$C_i(t) = \int_{v_i}^{v_{i+1}} c(v, t) dv \quad (4.6)$$

Eq. (4.4) can be recast in terms of the molar concentration by integrating over the i -th interval with substitution by Eq. (4.5), giving the following discretized ODE:

$$\frac{dC_i(t)}{dt} = B^\gamma - D^\gamma \quad (4.7)$$

where B^γ and D^γ are the birth and death terms respectively:

$$\begin{aligned}
B^\gamma &= \int_{v_i}^{v_{i+1}} \int_v^\infty k^\gamma(w) \delta(v - [w - v_m]) \sum_{j=2}^\infty C_j(t) \delta(w - x_j) dw dv \\
&= \int_{x_i}^{x_{i+1}} K(v, x_i) \int_v^\infty k^\gamma(w) \delta(v - [w - v_m]) \sum_{j=2}^\infty C_j(t) \delta(w - x_j) dw dv \\
&\quad + \int_{x_{i-1}}^{x_i} L(v, x_i) \int_v^\infty k^\gamma(w) \delta(v - [w - v_m]) \sum_{j=2}^\infty C_j(t) \delta(w - x_j) dw dv
\end{aligned} \tag{4.8}$$

$$D^\gamma = \int_{v_i}^{v_{i+1}} k^\gamma(v) \sum_{j=2}^\infty C_j(t) \delta(v - x_j) dv = k_i^\gamma C_i(t) \tag{4.9}$$

The partitioning of the birth integral from $[v_i, v_{i+1}]$ to $[x_i, x_{i+1}]$ and $[x_{i-1}, x_i]$ in Eq. (4.8) is the core of the FP technique, cf. S. Kumar and Ramkrishna (1996a). The fractions of polymers assigned to x_i from those that fall between $[x_i, x_{i+1}]$ and $[x_{i-1}, x_i]$, i.e. $K(v, x_i)$ and $L(v, x_i)$ respectively, are derived to ensure that the resulting equations are internally consistent with regard to the zeroth and the first moments (cf. Section A.1 of Appendix A). These are given as:

$$K(v, x_i) = \frac{x_{i+1} - v}{x_{i+1} - x_i} \tag{4.10}$$

$$L(v, x_i) = \frac{v - x_{i-1}}{x_i - x_{i-1}} \tag{4.11}$$

With these, Eqs. (4.7) – (4.9) simplify to:

$$\frac{dC_i(t)}{dt} = \sum_{j=i}^\infty n_{ij}^\gamma k_j^\gamma C_j(t) - k_i^\gamma C_i(t) \tag{4.12}$$

where n_{ij}^γ which represents the fractional allocation of polymers splitting from $DP = j$ into i is given as:

$$n_{ij}^\gamma = \int_{x_i}^{x_{i+1}} \left[\frac{x_{i+1} - v}{x_{i+1} - x_i} \right] \delta(v - [x_j - v_m]) dv + \int_{x_{i-1}}^{x_i} \left[\frac{v - x_{i-1}}{x_i - x_{i-1}} \right] \delta(v - [x_j - v_m]) dv \tag{4.13}$$

As such, $j \geq i$ is a basic requirement. Here, $v_m = 1$ is the DP of the monomer. Equations (4.12) – (4.13) form the foundation of Section 4.1.3.

4.1.3 Meshing and Implementation

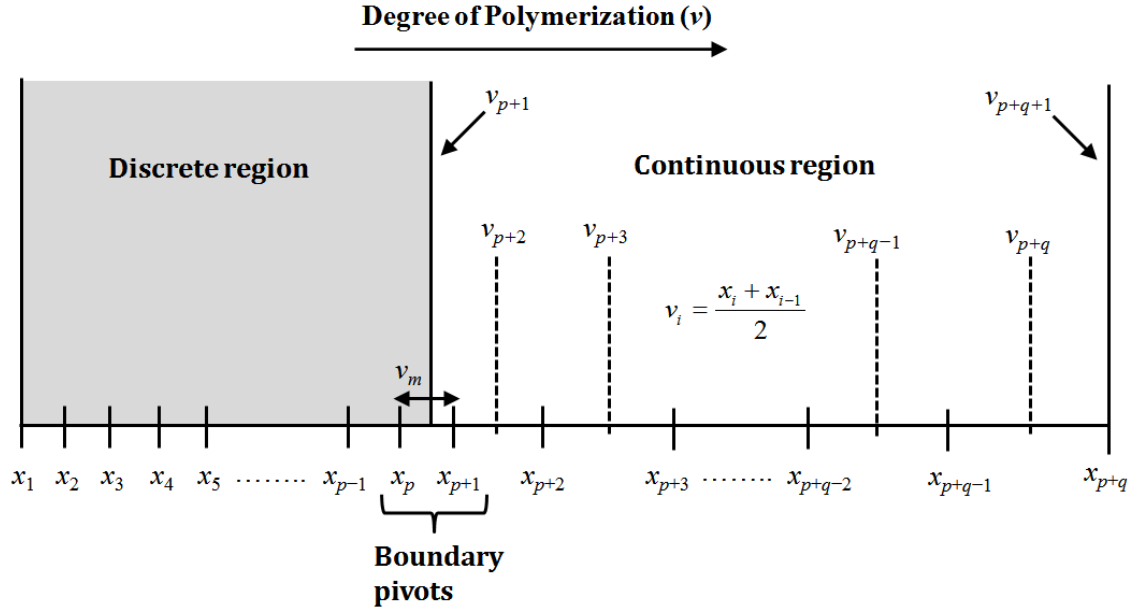


Figure 4.1: The discrete-continuous strategy with x_i being the pivot encompassed by v_i and v_{i+1} . The integer p is the number of pivots in the discrete region while q is the number of pivots in the continuous region, with $x_{p+q} = v_{p+q+1} = N$ ($N = \text{maximum DP}$).

To formulate the proposed FP solution for chain-end scission, by taking a cue from Gelbard and Seinfeld (1979), S. Kumar and Ramkrishna (1996a), and Kostoglou (2000), the DP axis is divided into two regions: the lower region has pivots placed v_m apart, while the higher region has pivots spaced wider than v_m . The latter makes physical sense as v_m is the smallest scission unit. Essentially it is a discrete-continuous split of the domain. This will allow the temporal evolution of small oligomers to be captured exactly, while economizing on the number of equations to solve for at the higher DPs. As illustrated in Figure 4.1, the number of pivots for the discrete region is p while the number of pivots in the continuous region is q (with $x_{p+q} = v_{p+q+1} = N$). The value of p (usually small compared to N) can be chosen such that it fulfils the requirements of the application considered. The continuous region treats the DP as a continuous variable from v_{p+1} to $v_{p+q+1} (= N)$. The continuous DP is then divided into q DP intervals. Throughout this thesis, a geometric

mesh $\left[x_{i+1} = x_i + r(x_i - x_{i-1}) \right]$ was employed in the continuous region where $r = \left[x_{p+q} / x_{p+1} \right]^{\frac{1}{q-1}}$ is the geometric ratio for q pivot points between x_{p+1} and x_{p+q} . Following S. Kumar and Ramkrishna (1996a), here $v_i = [x_i + x_{i-1}] / 2$ is used. Unlike the discrete region, there are no restrictions in the position of the pivots in the continuous region. This feature of the continuous region allows a coarser mesh to be deployed ($p+q \ll N$) and therefore reduces the number of ODEs to be solved.

By requiring that the smallest gap in the continuous region to exceed v_m , i.e. $x_{p+2} - x_{p+1} > v_m$, it follows that $rx_{p+1} - x_{p+1} > v_m$ leading to the following inequality:

$$r > 1 + v_m / x_{p+1} \quad (4.14)$$

Substituting $r = \left[x_{p+q} / x_{p+1} \right]^{\frac{1}{q-1}}$ into Eq. (4.14) and simplifying further by taking the natural logarithm on both sides of the equation yields:

$$q < 1 + \left\{ \ln \left(\frac{x_{p+q}}{x_{p+1}} \right) / \ln \left(1 + \frac{v_m}{x_{p+1}} \right) \right\} \quad (4.15)$$

Using the relationships of $x_{p+q} = N$ and $x_{p+1} = p + v_m$, Eq. (4.15) can be rewritten as:

$$q < 1 + \left\{ \ln \left[\frac{N}{p + v_m} \right] / \ln \left[1 + \frac{v_m}{p + v_m} \right] \right\} \quad (4.16)$$

Equation (4.16) is a useful analytical expression to assist the proper choice of the value of q , given those for p , N and v_m .

To obtain the final set of equations to be solved, the main task is to obtain Eqs. (4.12) – (4.13) for both the discrete and the continuous regions. Proper connection of both regions requires that mass is conserved in the final set of equations. For the case $i = p+1$, the corresponding $n_{p+1,j}^\gamma$ is given as:

$$n_{p+1,j}^{\gamma} = \int_{x_{p+1}}^{x_{p+2}} \left[\frac{x_{p+2} - v}{x_{p+2} - x_{p+1}} \right] \delta(v - [x_j - v_m]) dv + \int_{x_p}^{x_{p+1}} \left[\frac{v - x_p}{x_{p+1} - x_p} \right] \delta(v - [x_j - v_m]) dv \quad (4.17)$$

From Eq. (4.17), at $j = p+1$, the first term on the RHS vanishes because the argument of the Dirac delta function $v = x_{p+1} - v_m = x_p$ does not fall within the interval of integration $[x_{p+1}, x_{p+2}]$. The second integral on the RHS, on the other hand, has $v = x_{p+1} - v_m = x_p$ which lies on the lower limit of the integral. This integral can be recast into the following form:

$$\int_a^b f(v) \delta(v - [a + \chi]) dv; \quad a = x_p; \quad b = x_{p+1}; \quad f(v) = \frac{v - x_p}{x_{p+1} - x_p} \quad (4.18)$$

Given that $0 < \chi < b - a$, the evaluation of Eq. (4.18) is simply $f(a + \chi)$. When expanded using the Taylor series at $v = a$, this results in:

$$f(a + \chi) = f(a) + \chi \frac{d}{dv} f(a) + \frac{\chi^2}{2!} \frac{d^2}{dv^2} f(a) + \dots \quad (4.19)$$

By using a limiting process, it follows that:

$$\begin{aligned} \lim_{\chi \rightarrow 0} \int_a^b f(v) \delta[v - (a + \chi)] dv &= \lim_{\chi \rightarrow 0} f(a + \chi) \\ &\approx \lim_{\chi \rightarrow 0} \left[f(a) + \chi \frac{d}{dv} f(a) + \frac{\chi^2}{2!} \frac{d^2}{dv^2} f(a) + \dots \right] = f(a) \end{aligned} \quad (4.20)$$

Using the results of Eq. (4.20), the second term on the RHS of Eq. (4.17) at $j = p+1$ also vanishes, hence $n_{p+1,p+1}^{\gamma} = 0$. Moving forward to $j = p+2$, the second term on the RHS of Eq. (4.17) vanishes because $v = x_{p+2} - v_m > x_{p+1}$ does not fall within $[x_p, x_{p+1}]$. For the first

term, $v = x_{p+2} - v_m$ falls within $[x_{p+1}, x_{p+2}]$, thus $n_{p+1,p+2}^{\gamma} = \frac{v_m}{x_{p+2} - x_{p+1}}$. Beyond $j = p+2$,

$n_{p+1,j}^{\gamma} = 0$ because $x_j - v_m$ does not fall within either interval of integration.

Using similar arguments, for $i \geq p+2$, Eq. (4.13) can be evaluated easily as above because the distance between adjoining pivots is greater than v_m . For these cases, the only

two non-zero entries are $n_{ii}^\gamma = 1 - v_m / (x_i - x_{i-1})$ and $n_{i,i+1}^\gamma = v_m / (x_{i+1} - x_i)$. It can be said that the case for $i = p+1$ at the discrete-continuous border is special because $n_{p+1,p+1}^\gamma = 0$. Further, at the last pivot $i = p+q$, clearly the term $n_{p+q,p+q+1}^\gamma$ is inapplicable.

For $i = 2$ to p , there are cases where $\delta(v - [x_j - v_m])$ is non-zero at the upper limit of the integral. To illustrate, consider Eq. (4.21) below for $i = p$:

$$n_{pj}^\gamma = \int_{x_p}^{x_{p+1}} \left[\frac{x_{p+1} - v}{x_{p+1} - x_p} \right] \delta(v - [x_j - v_m]) dv + \int_{x_{p-1}}^{x_p} \left[\frac{v - x_{p-1}}{x_p - x_{p-1}} \right] \delta(v - [x_j - v_m]) dv \quad (4.21)$$

At $j = p+1$, the first term on the RHS equals unity using the limiting process described above. For the second term, $v = x_{p+1} - v_m = x_p$ lies on the upper limit of the integral. Because the limiting process assumes that $(x_p + \chi)$ falls within $[x_p, x_{p+1}]$, therefore the same $(x_p + \chi)$ must lie outside $[x_{p-1}, x_p]$, which causes the second term to vanish. For clarity, the following convention is adopted:

$$\int_b^a f(v) \delta(v - b) dv = f(b), \quad \int_c^b f(v) \delta(v - b) dv = 0, \quad c < b < a \quad (4.22)$$

Using this, the limiting process applied to the Dirac delta function yields identically the discrete rate equations. Collating all cases for $i = 1$ to $p+q$, the following general set of equations is obtained:

Discrete region:

$$\frac{dC_1(t)}{dt} = 2k_2^\gamma C_2(t) + \sum_{j=3}^{p+q} k_j^\gamma C_j(t) \quad (4.23)$$

$$\frac{dC_i(t)}{dt} = k_{i+1}^\gamma C_{i+1}(t) - k_i^\gamma C_i(t) \quad i = 2, 3, \dots, p-1 \quad (4.24)$$

Discrete-continuous boundary:

$$\frac{dC_p(t)}{dt} = k_{p+1}^\gamma C_{p+1}(t) - k_p^\gamma C_p(t) \quad (4.25)$$

$$\frac{dC_{p+1}(t)}{dt} = \left[\frac{v_m}{x_{p+2} - x_{p+1}} \right] k_{p+2}^\gamma C_{p+2}(t) - k_{p+1}^\gamma C_{p+1}(t) \quad (4.26)$$

Continuous region:

$$\frac{dC_i(t)}{dt} = \left[\frac{v_m}{x_{i+1} - x_i} \right] k_{i+1}^\gamma C_{i+1}(t) + \left[1 - \frac{v_m}{x_i - x_{i-1}} \right] k_i^\gamma C_i(t) - k_i^\gamma C_i(t) \quad (4.27)$$

$$i = p+2, p+3, \dots, p+q-1$$

$$\frac{dC_{p+q}(t)}{dt} = \left[1 - \frac{v_m}{x_{p+q} - x_{p+q-1}} \right] k_{p+q}^\gamma C_{p+q}(t) - k_{p+q}^\gamma C_{p+q}(t) \quad (4.28)$$

Equations (4.23) – (4.28) describe chain-end scission and the system of equations constitutes $p+q$ ODEs which is less than N ODEs in the fully discrete solution.

In this formulation, the molar concentration profiles $C_i(t)$ of the experimentally detectable oligomers are solved for directly. Where the molar concentration density c_i is required instead, by applying the mean value theorem on $C_i(t) = \int_{v_i}^{v_{i+1}} c(v, t) dv$, it can be shown that $c_i = C_i / [v_{i+1} - v_i]$. In the discrete region, as $v_{i+1} - v_i = 1$, $c_i = C_i$. Section 4.1.4 will be devoted to validating the performance of the FP solution against the exact solutions given by Eqs. (4.1) – (4.3) in Section 4.1.1.

4.1.4 Case Study on Chain-End Scission

To validate the performance of the proposed solution approach, chain-end scission of a polymer with a broad distribution and a large N was studied. Studies of such a system are important particularly for the depolymerization of polymers, e.g. the enzymatic hydrolysis of starch where glucoamylase removes successive glucose units from the non-reducing end (Breuninger et al., 2009). Here, the polymer was assumed to be starch with the monomer glucose (molecular weight = 180 g/mol) and the subsequent oligomers

having molecular weights of 162DP+18 g/mol. As the main objective here is to elucidate the efficacy of the FP technique in solving chain-end scission rather than to very realistically mimic the intricate biochemical reactions, only linear polymers with simple scission kinetics were considered in this illustrative example. As had been adopted in the work of Chang et al. (2002), the initial molar concentration density was assumed to follow the Schulz-Zimm distribution (Dole, 1972), as given in Eqs. (3.1) - (3.3) in Section 3.3.1. The number-average DP (\overline{M}_n) and the weight-average DP (\overline{M}_w) required in these equations can be obtained as:

$$\overline{M}_n(t) = \frac{\text{1st moment}}{\text{0th moment}} = \frac{\int_0^\infty v c(v, t) dv}{\int_0^\infty c(v, t) dv} = \frac{\sum_{i=1}^\infty x_i C_i}{\sum_{i=1}^\infty C_i} \quad (4.29)$$

$$\overline{M}_w(t) = \frac{\text{2nd moment}}{\text{1st moment}} = \frac{\int_0^\infty v^2 c(v, t) dv}{\int_0^\infty v c(v, t) dv} = \frac{\sum_{i=1}^\infty x_i^2 C_i}{\sum_{i=1}^\infty x_i C_i} \quad (4.30)$$

In this study, values of the parameters for sweet potato starch (Breuninger et al., 2009) were used, as in Table 4.1. To calculate the initial conditions ($C_i(0) = \int_{v_i}^{v_{i+1}} c(v, 0) dv$) which involve integration along the v axis, meshing had to be done on the v axis for both the exact and the FP solutions. For this purpose, to obtain the limits of the integral for the exact formulation, let y be the discrete DPs for the exact formulation: $y_1 = 1, y_2 = 2, \dots, y_{N-1} = N-1, y_N = N$. Following this, $v_1 = y_1$ and $v_{N+1} = y_N$, with those in between calculated by $v_i = [y_i + y_{i-1}]/2$. The FP formulation adopts the same procedure for computing the limits of the integral, with $v_1 = x_1 = 1, v_{p+q+1} = x_{p+q} = N$ and $v_i = [x_i + x_{i-1}]/2$. Geometric meshing was used in the continuous region for the FP formulation, i.e. $[x_{p+1}, x_{p+q}]$.

Table 4.1: Values of parameters (Breuninger et al., 2009) used in the case study on chain-end scission

Parameter	Value	Units
Maximum DP (N)	22496	-
Initial mass concentration of the polymer ($m_S(0)$)	10	g/L
Number-average DP (\overline{M}_n)	4100	-
Weight-average DP (\overline{M}_w)	5430	-

In the simulations presented in this section, the DP-dependent rate kernel (S. Kumar & Ramkrishna, 1996a; Staggs, 2004), i.e. $k^\gamma(v) = k_p^\gamma v$, was adopted in both the exact and the FP solution, where $k_p^\gamma = 1/80$. The arbitrary value was chosen mainly to have a total evolution time of around 1000 time units. In the results that follow, this time span was normalized against the time required for 99% monomer production ($t_{99\%}$), i.e. the dimensionless time $\theta = t/t_{99\%}$. (cf. Section 3.3.3) The total number of equations for the FP solution was $p+q$ whereas for the exact solution, N equations were solved. In the FP solution, values of p and q are the remaining degrees of freedom. Values of p and q were chosen to ensure that the smallest gap in the continuous region exceeds v_m according to Eq. (4.16). Unless mentioned otherwise, to demonstrate the efficacy of the solution approach, values of $[p, q] = [100, 500]$ which totalled less than 3% of the N equations used in the exact solution were chosen below. Integration of the ODEs for both the exact and the FP formulation in Figures 4.2 – 4.7 was done using the ‘ode45’ subroutine in MATLAB® R2009b on the workstation A (cf. Section 3.1 for the specifications of the workstation). Example codes for the simulation of chain-end scission using the FP and the exact solution are given in Sections A.2 – A.3 of Appendix A.

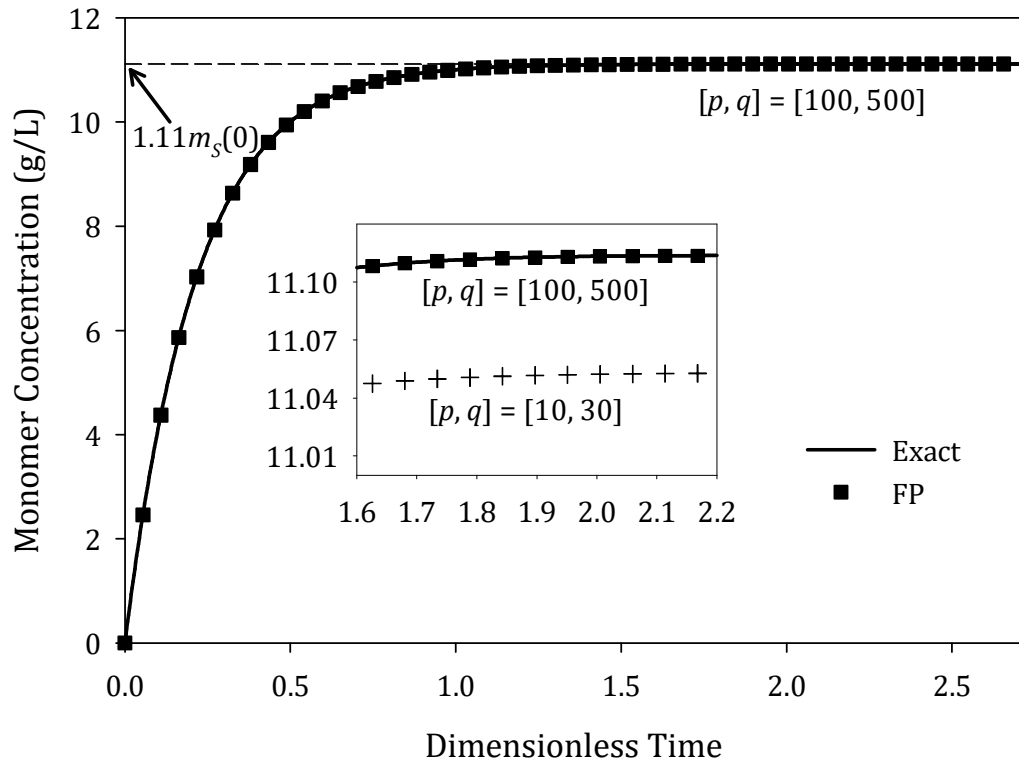


Figure 4.2: Transient of the mass concentration of monomer (glucose) using the exact and fixed pivot (FP) solutions for chain-end scission. Here $[p, q] = [100, 500]$, $r = 1.0109$ and $m_s(0) = 10$ g/L. The dimensionless time is normalized against the time required for 99% monomer production. The error in the initial mass due to discretization $[\epsilon_s, \text{cf. Eq. (3.9)}]$ is $2.61 \times 10^{-4} m_s(0)$. The case of $[p, q] = [10, 30]$ with $r = 1.3007$ was given as a reference to coarsely resolved mesh.

Figure 4.2 shows the transient for the mass concentration of monomer glucose using the exact and FP solutions. Theoretically, if all starch is hydrolyzed to produce glucose, the final mass concentration of glucose is $1.11m_s(0)$ ($m_s(0) = 10$ g/L). This was indeed observed. Additionally, the mass concentration of the monomer was rather insensitive to the total number of pivots. Even with a very coarse mesh, i.e. $[p, q] = [10, 30]$, ϵ_g as defined by Eq. (3.8) is $\sim O(10^{-3})$ whereas an increase in the total number of pivots from 200 to 600 (not shown) yielded almost identically accurate results, with ϵ_g ranging from $\sim O(10^{-4})$ to $\sim O(10^{-6})$. This rather insensitive concentration profile of the monomer to the

total number of pivots can be rationalized because the monomer equation is a mere collection of all the breakage terms from the larger polymers.

Figure 4.3 shows the concentration profiles of the DP2 – DP7 oligomers while Figure 4.4 shows the temporal evolution of the molar concentration density at different dimensionless times. The results showed that the transients of the mass concentration of oligomers were predicted satisfactorily with reference to the exact solutions, with $\epsilon_g = 3.9 \times 10^{-6}$ (for DP2) and $\epsilon_g = 6.7 \times 10^{-6}$ (for DP7) respectively. Since the un-normalized concentrations of these oligomers were of $\sim O(10^{-3})$ to $\sim O(10^{-4})$, ϵ_g which was of $\sim O(10^{-6})$ is satisfactory. The availability of the results in Figure 4.3 complements experimental studies because the concentrations of these compounds could be analyzed (Paolucci-Jeanjean, Belleville, Rios, & Zakhia, 2000; Paolucci-Jeanjean, Belleville, Zakhia, & Rios, 2000; Steverson, Korus, Admassu, & Heimsch, 1984). Similar performance was also observed in the molar concentration density, where even at a dimensionless time beyond 99% of monomer production ($\theta > 1$), good predictions were observed.

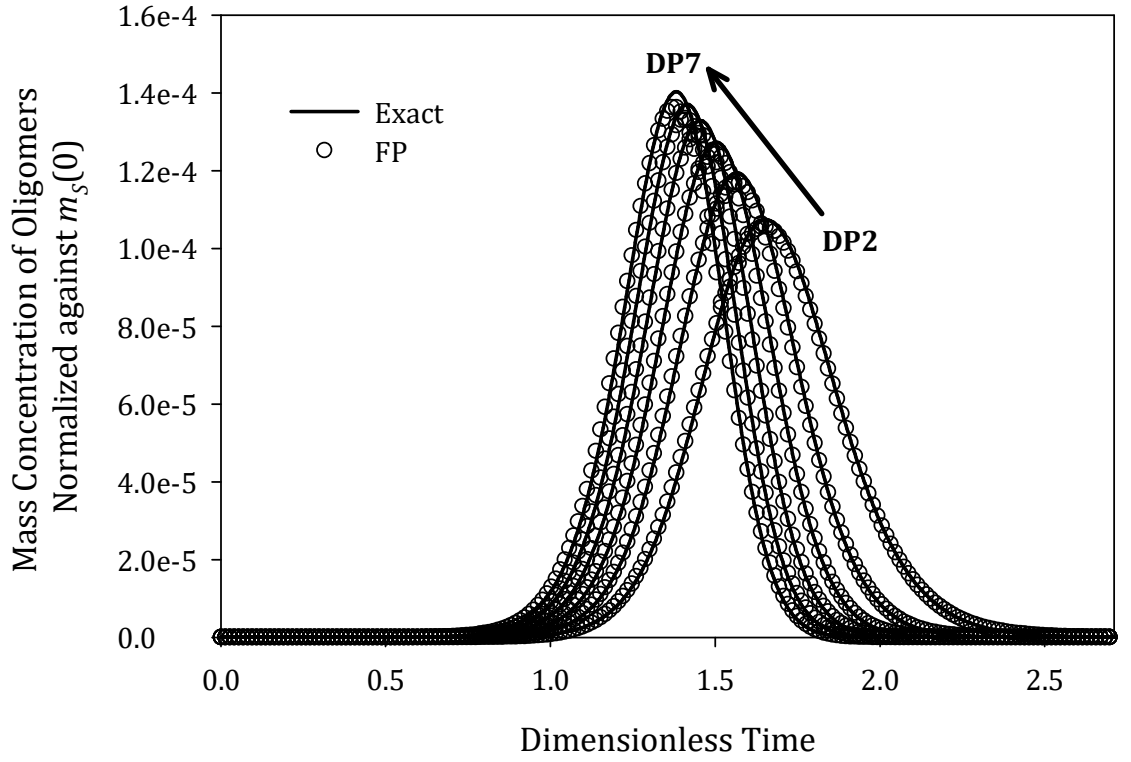


Figure 4.3: Transients of the normalized mass concentration of oligomers using the exact and the fixed pivot (FP) solutions for chain-end scission. The highest peak corresponds to the DP7 oligomer, followed by the DP6 – DP2 oligomers in the order of decreasing peak heights. Here $[p, q] = [100, 500]$, $r = 1.0109$, $m_s(0) = 10$ g/L, and the error in the initial mass due to discretization (ε_s) is $2.61 \times 10^{-4} m_s(0)$. The dimensionless time is normalized against the time required for 99% monomer production.

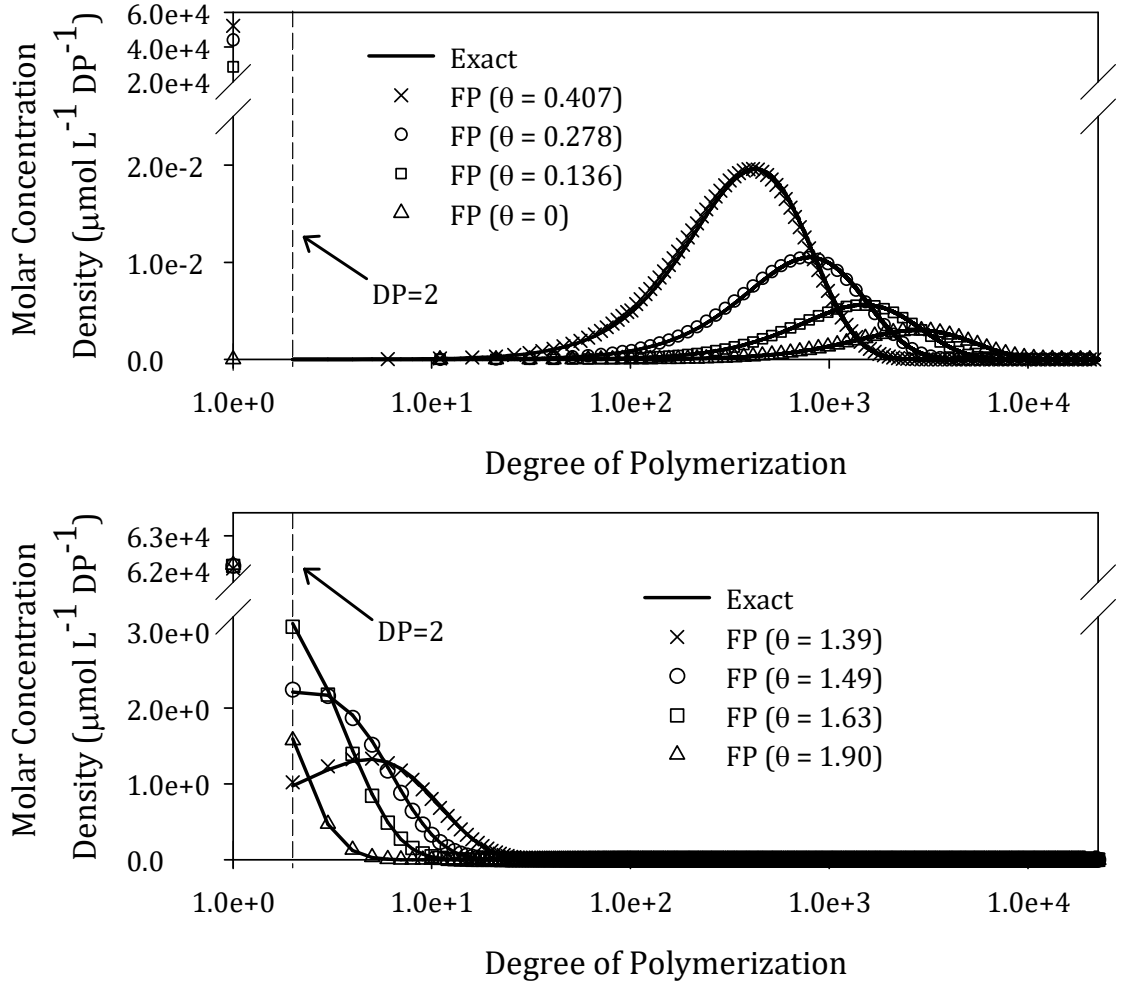


Figure 4.4: Transients of the molar concentration density using the exact and the fixed pivot (FP) solutions for chain-end scission. The dimensionless time (θ) is normalized against the time required for 99% monomer production. Here $[p, q] = [100, 500]$ and $r = 1.0109$. The exact solutions fall on the FP solutions for $DP = 1$.

Often in applications involving synthetic polymers, a more macroscopic measure instead is required. Examples of such measure are the number-average $[\overline{M}_n, \text{Eq. (4.29)}]$ and weight-average $[\overline{M}_w, \text{Eq. (4.30)}]$ DPs as well as the polydispersity index (PD), which can be computed from the moments of the polymer distribution as follows:

$$PD(t) = \frac{\overline{M}_w(t)}{\overline{M}_n(t)} \quad (4.31)$$

Figures 4.5 – 4.7 show the evolution of \overline{M}_n , \overline{M}_w and PD respectively using $[p, q] = [100, 500]$. As the FP technique was derived based on the principle of conserving moments, its capability in predicting moment-related quantities was also assessed at a very coarse mesh of $[p, q] = [10, 30]$. From the results, the prediction of \overline{M}_n was fairly accurate even at a very coarse mesh. Such a remarkable accuracy in predicting \overline{M}_n is not surprising, considering that it is computed solely from the zeroth and the first moments of the distribution, of which clearly the FP technique is capable of conserving. The prediction accuracy for \overline{M}_w , however, was more sensitive to the resolution of the mesh, with $\epsilon_g \sim O(10^{-1})$ at $[p, q] = [10, 30]$. As the mesh was refined further, $\epsilon_g \sim O(10^{-3})$ at $[p, q] = [100, 500]$. This observation is understandable because the computation of \overline{M}_w requires the second moment of the distribution, which was not chosen when deriving Eq. (4.13). The prediction accuracy for PD followed the same trend as that for \overline{M}_w , as the computation of the PD involves the use of the zeroth, the first, and the second moments of the polymer distribution. For PD, ϵ_g is $\sim O(10^{-3})$ at $[p, q] = [100, 500]$ and ϵ_g is $\sim O(10^{-1})$ at $[p, q] = [10, 30]$. Nonetheless, unless at the extreme case of $[p, q] = [10, 30]$ which totalled less than 0.2% of N , the ability to predict quantities requiring the second moment using merely less than 3% of N at $[p, q] = [100, 500]$ to a good degree of accuracy is remarkable considering that the formulation conserves only the zeroth and the first moments.

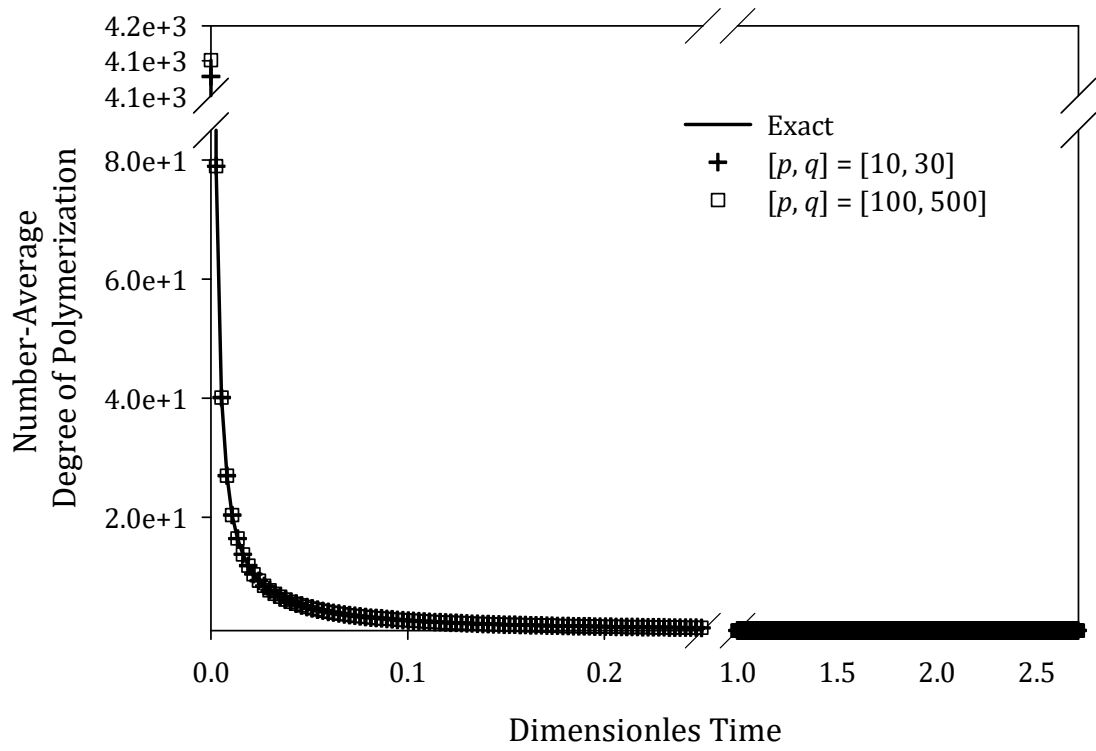


Figure 4.5: Transient of the number-average DP using the exact and the fixed pivot (FP) solutions for chain-end scission. Here, $r = 1.0109$ for $[p, q] = [100, 500]$. The case of $[p, q] = [10, 30]$ with $r = 1.3007$ was given as a reference to coarsely resolved mesh. The dimensionless time is normalized against the time required for 99% monomer production.

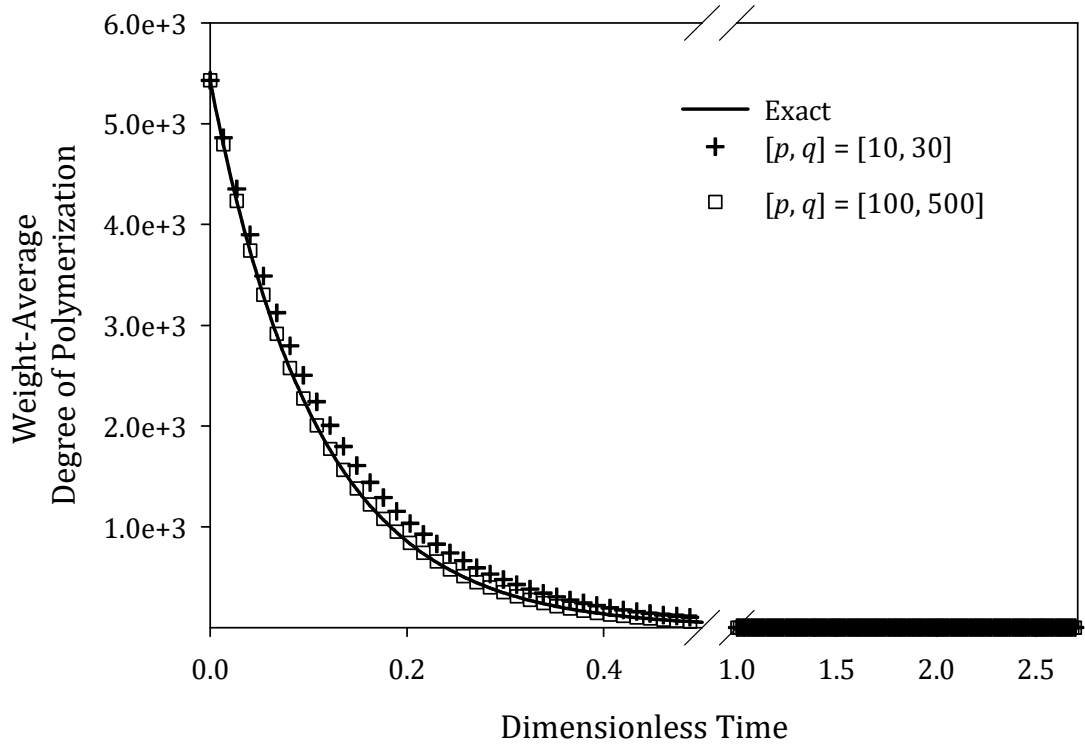


Figure 4.6: Transient of the weight-average DP using the exact and the fixed pivot (FP) solutions for chain-end scission. Here, $r = 1.0109$ for $[p, q] = [100, 500]$. The case of $[p, q] = [10, 30]$ with $r = 1.3007$ was given as a reference to coarsely resolved mesh. The dimensionless time is normalized against the time required for 99% monomer production.

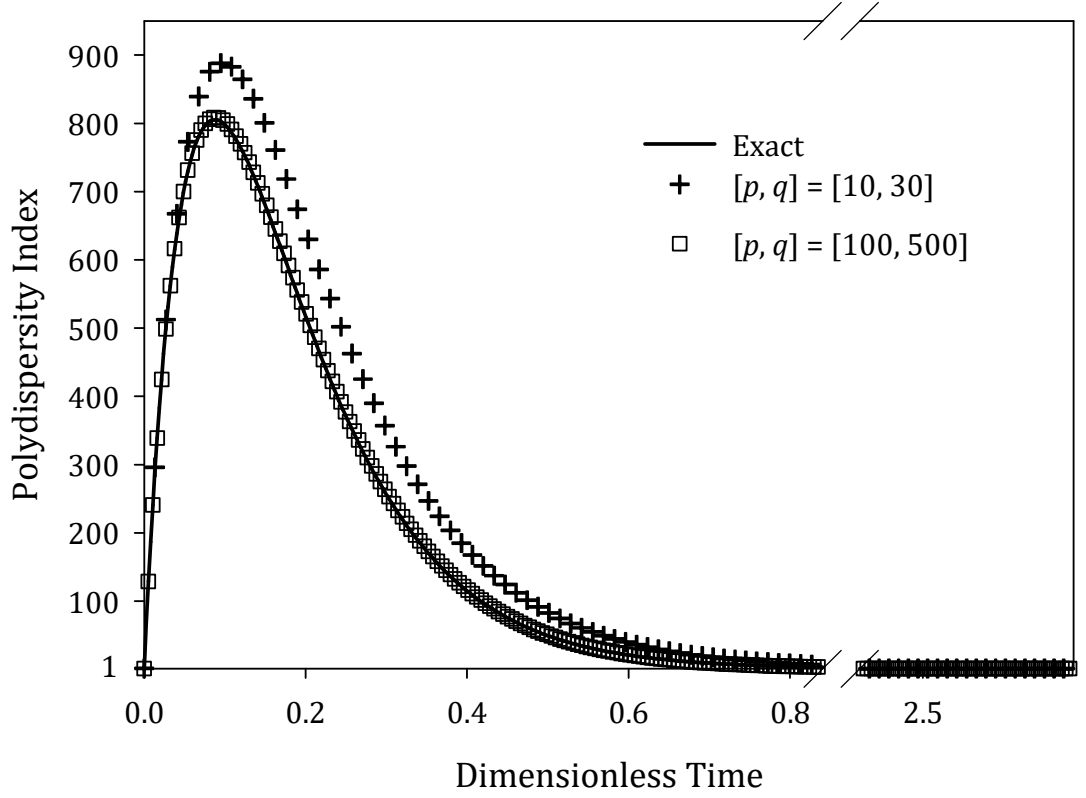


Figure 4.7: Transient of the polydispersity index using the exact and the fixed pivot (FP) solutions for chain-end scission. Here $[p, q] = [100, 500]$ and $r = 1.0109$. The case of $[p, q] = [10, 30]$ with $r = 1.3007$ was given as a reference to coarsely resolved mesh. The dimensionless time is normalized against the time required for 99% monomer production.

Due to the drastic reduction in the number of ODEs to solve, the computation time required to produce the FP solution is much less than that required to solve the exact case. The benefit is even more evident when the size of the problem increases. To illustrate, the simulations were re-run after scaling up the parameters in Table 4.1, i.e. N , \overline{M}_n , and \overline{M}_w , by identical factors while retaining the value of $m_s(0)$. The values of \overline{M}_n and \overline{M}_w were scaled together with N to maintain the PD of the distribution. For every set of N , \overline{M}_n , and \overline{M}_w tested, five different $p+q$'s were studied, i.e. $[p, q] = [100, 500]$, $[p, q] = [150, 650]$, $[p, q] = [200, 800]$, $[p, q] = [250, 950]$, and $[p, q] = [300, 1200]$, all of which obeyed the constraint of Eq. (4.16). Choosing a very small value of p (e.g. $p = 10$) would

require a corresponding drastic reduction in the value of q to fulfil the constraint, and a consequent loss in fidelity in capturing the evolution of the continuous region. For approximately $p \sim O(10^2)$, such loss of accuracy was no longer an issue. The computational time for the FP and the exact solutions achieved using workstation B (cf. Section 3.1 for the specifications of the workstation) were $t_{FP}(s) = 3.080 \times [(p+q)/10^3]^{1.96}$ and $t_{Exact}(s) = 1.403 \times [N/10^3]^{1.96}$ respectively for the range of DP studied. Consolidating this information, Figure 4.8 shows the fraction of computation time required by the FP solution as a function of $(p+q)/N$. A power law expression of $t_{FP}/t_{Exact} = 2.195 \times [(p+q)/N]^{1.96}$ was obtained. The scatter was mainly due to variability clocked by repeated runs of the FP solutions, all of which completed within seconds or less. It is noteworthy that even at the largest $p+q$ tested, the computational time required by the FP technique was still below 1% of that required by the exact simulation.

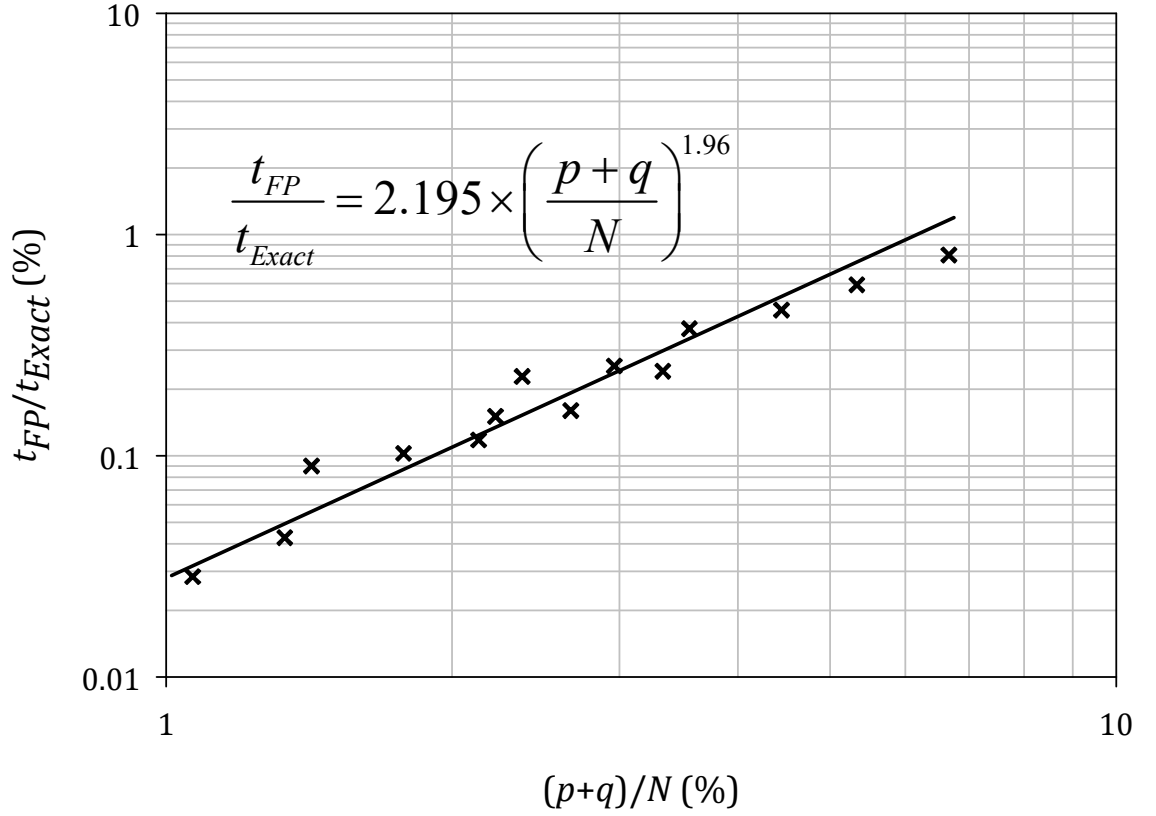


Figure 4.8: Fraction of time required to produce the fixed pivot solution vs. the fraction of pivots used. Here, t_{FP} is the time required to obtain the fixed pivot solution, t_{Exact} is the time required to obtain the exact solution.

To further showcase the efficacy of the FP technique, the resulting molar concentration densities at $\theta = 1.08$ for $[p, q] = [100, 500]$ (i.e. the smallest fraction of pivots tested) as the value of N increased are shown in Figure 4.9. At fractions of pivots below 3% of N , the predictions remained satisfactory. Over-predictions, however, were observed where the molar concentration density varied steeply. This is an inherent nature of the FP technique, as expounded by S. Kumar and Ramkrishna (1996a). To solve even larger problems, the accuracy of the prediction can be increased by mesh refinement, i.e. using more pivots. For $N = 10 \times 22496$, the number of pivots was slightly increased to $p+q = 2000$ (still $< 1\%$ of N). The result remained very satisfactory, as illustrated by Figure 4.10. On the workstation B, the computational time required by the solution technique was

merely 0.02 % (≈ 9.4 s) of the duration (≈ 16 hrs) required by the exact case. Thus, the benefits of the FP technique are evident when solving large problems.

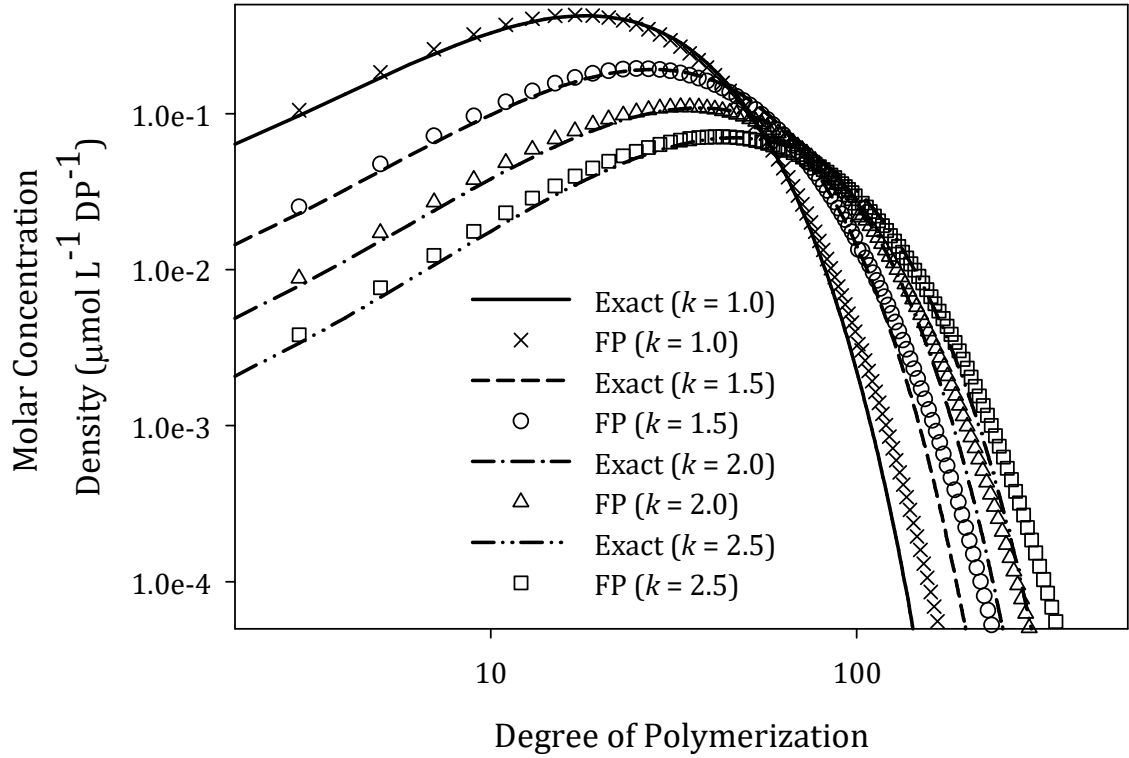


Figure 4.9: Effect of increasing the size of the exact problem according to $N = k \times 22496$ ($k = 1.0, 1.5, 2.0, 2.5$) while retaining the initial polydispersity index at 1.32 as well as using $[p, q] = [100, 500]$ for the fixed pivot (FP) solution. The geometric ratios employed were $r = 1.0109, 1.0117, 1.0123$, and 1.0128 in the order of increasing k . The molar concentration density is shown for the dimensionless time, $\theta = 1.08$. Here, $m_S(0) = 10$ g/L and the errors in the initial mass due to discretization (ϵ_S) were $2.6 \times 10^{-4} m_S(0)$, $1.7 \times 10^{-4} m_S(0)$, $1.2 \times 10^{-4} m_S(0)$, and $0.95 \times 10^{-4} m_S(0)$ in the increasing order of k .

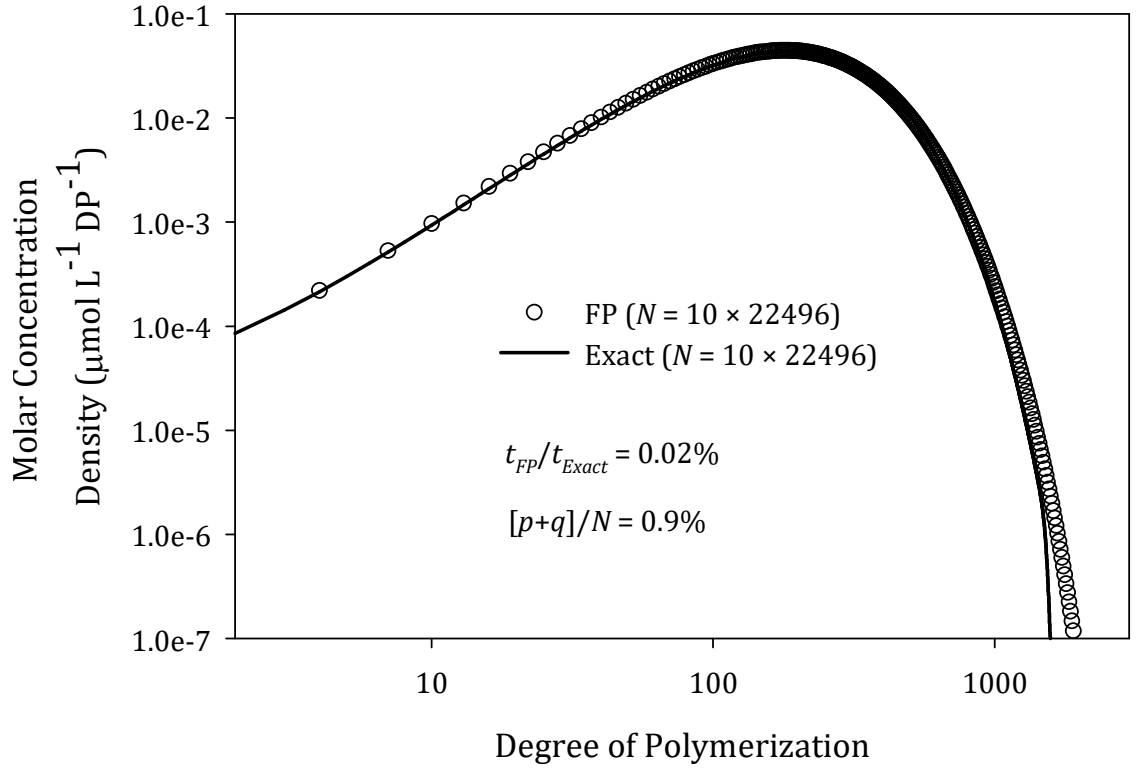


Figure 4.10: Performance of the fixed pivot (FP) solution in solving a relatively large chain-end scission problem, i.e. $N = 224,960$, by using $< 1\%$ of the number of equations employed by the exact solution. Here, the initial polydispersity index was 1.32, $[p, q] = [256, 1744]$, $r = 1.0039$, $m_S(0) = 100$ g/L and the error in the initial mass due to discretization (ε_S) was $0.26 \times 10^{-4} m_S(0)$. The molar concentration density is shown for the dimensionless time, $\theta = 1.08$ when slightly more than 99% of the monomers had been formed. Here t_{FP} and t_{Exact} are respectively the time required to produce the FP and the exact solutions.

4.1.5 Observations of the FP Method for Solving Chain-End Scission

From Eq. (4.27), the rate of change in the concentration of the polymer at the i -th pivot can be conveniently simplified as:

$$\frac{dC_i(t)}{dt} = v_m \left[\frac{k_{i+1}^\gamma C_{i+1}(t)}{x_{i+1} - x_i} - \frac{k_i^\gamma C_i(t)}{x_i - x_{i-1}} \right] \quad (4.32)$$

In general, if geometric meshing is used, $r = \frac{x_i - x_{i-1}}{x_{i-1} - x_{i-2}} = \frac{v_i - v_{i-1}}{v_{i-1} - v_{i-2}}$. Using this

relationship and rewriting Eq. (4.32) in terms of c_i (the average molar concentration density) as shown below, a more compact expression results in Eq. (4.34):

$$\begin{aligned} \frac{d}{dt}[c_i(t)(v_{i+1} - v_i)] &= v_m \left\{ \frac{k_{i+1}^\gamma [c_{i+1}(t)(v_{i+2} - v_{i+1})]}{x_{i+1} - x_i} - \frac{k_i^\gamma [c_i(t)(v_{i+1} - v_i)]}{x_i - x_{i-1}} \right\} \\ &= v_m \left\{ \frac{k_{i+1}^\gamma [c_{i+1}(t)r(v_{i+1} - v_i)]}{x_{i+1} - x_i} - \frac{k_i^\gamma [c_i(t)(v_{i+1} - v_i)]}{(1/r)(x_{i+1} - x_i)} \right\} \\ &= rv_m(v_{i+1} - v_i) \left\{ \frac{k_{i+1}^\gamma c_{i+1}(t) - k_i^\gamma c_i(t)}{x_{i+1} - x_i} \right\} \end{aligned} \quad (4.33)$$

$$\therefore \frac{dc_i(t)}{dt} = rv_m \left[\frac{k_{i+1}^\gamma c_{i+1}(t) - k_i^\gamma c_i(t)}{x_{i+1} - x_i} \right] \quad (4.34)$$

Four observations can be gained from Eq. (4.34). First, if $r = v_m = 1$, giving an arithmetic series with $x_{i+1} - x_i = 1$, the equation reduces as required to the exact solution, i.e.

$dc_i(t)/dt = k_{i+1}^\gamma c_{i+1}(t) - k_i^\gamma c_i(t)$. In the common case with $v_m = 1$, the departure of the

value of r from unity could then be interpreted as the extent of deviation from the exact formulation. Therefore, unlike other forms of depolymerization, e.g. continuous random scission, where values of r as large as 2 – 3 had been used (S. Kumar & Ramkrishna, 1996a), the same luxury cannot be afforded by the kernel of the chain-end scission.

Fortunately, values of $r = [x_{p+q}/x_{p+1}]^{\frac{1}{q-1}}$ being greater and close to 1 are frequently encountered because $q-1$ is usually large. For instance, $r = 1.0109$ for the results presented in Figures. 4.2 – 4.7 for $[p, q] = [100, 500]$.

Second, under a specific form of discretization, Eq. (4.34) corroborates the approximate continuous model for chain-end scission using the first order Taylor series expansion (Kostoglou, 2000):

$$\frac{\partial c(v, t)}{\partial t} = v_m \frac{\partial}{\partial v} [k^\gamma(v) c(v, t)] \quad (4.35)$$

Approximating Eq. (4.35) by the finite forward difference method and employing a change of variable from v to x via $v_i = [x_i + x_{i-1}]/2$ as shown in Eq. (4.36) yields Eq.

(4.37):

$$\begin{aligned} \frac{dc_i(t)}{dt} &= v_m \left\{ \frac{k_{i+1}^\gamma c_{i+1}(t) - k_i^\gamma c_i(t)}{v_{i+1} - v_i} \right\} \\ &= v_m \left\{ \frac{k_{i+1}^\gamma c_{i+1}(t) - k_i^\gamma c_i(t)}{(x_{i+1} + x_i)/2 - (x_i + x_{i-1})/2} \right\} \\ &= v_m \left\{ \frac{k_{i+1}^\gamma c_{i+1}(t) - k_i^\gamma c_i(t)}{[x_{i+1} + (1/r)x_{i+1}]/2 - [x_i + (1/r)x_i]/2} \right\} \end{aligned} \quad (4.36)$$

$$\therefore \frac{dc_i(t)}{dt} = v_m \left[\frac{k_{i+1}^\gamma c_{i+1}(t) - k_i^\gamma c_i(t)}{(1/2)(1+1/r)(x_{i+1} - x_i)} \right] \quad (4.37)$$

If $r = 1$, i.e. arithmetic meshing, Eq. (4.34) and Eq. (4.37) coincide. In other words, although derived from a different standpoint, the FP equations for chain-end scission can be solved with the same performance as that for Eq. (4.35) as discretized above.

The third observation pertains to the limitation of the FP technique in modelling chain-end scission. The fact that Eq. (4.34) is a two-point finite difference approximation to the slope poses accuracy problems when dealing with polymers which exhibit narrow distributions. To illustrate, the FP technique was implemented on the same example problem used in Stickel and Griggs (2012). In the example, chain-end scission was modelled using the continuous distribution kinetics on the initial distribution given by Eq. (3.4). The values of c_{in} , ω and κ were given as 100, 400 and 1.5 respectively. Integration of the ODEs for both the exact and the FP formulation was similarly done using the ‘ode45’ subroutine on the workstation A and the example codes for simulating the FP and exact solutions are given in Sections A.4 – A.5 of Appendix A. Figure 4.11 shows the results

of the simulation. In this case, arithmetic meshing in the continuous region ensures more closely spaced pivots to capture the sharp changes in the molecular distribution. A common difference (or distance between pivots) of 1.5 was used in the continuous region, resulting in $[p, q] = [10, 545]$. With such a fine mesh, i.e. $\approx 67\%$ of the $N = 827$ equations used for the exact case, the results showed that as time progressed, over-predictions were observed near the base of the peak whereas under-predictions were observed near the peak.

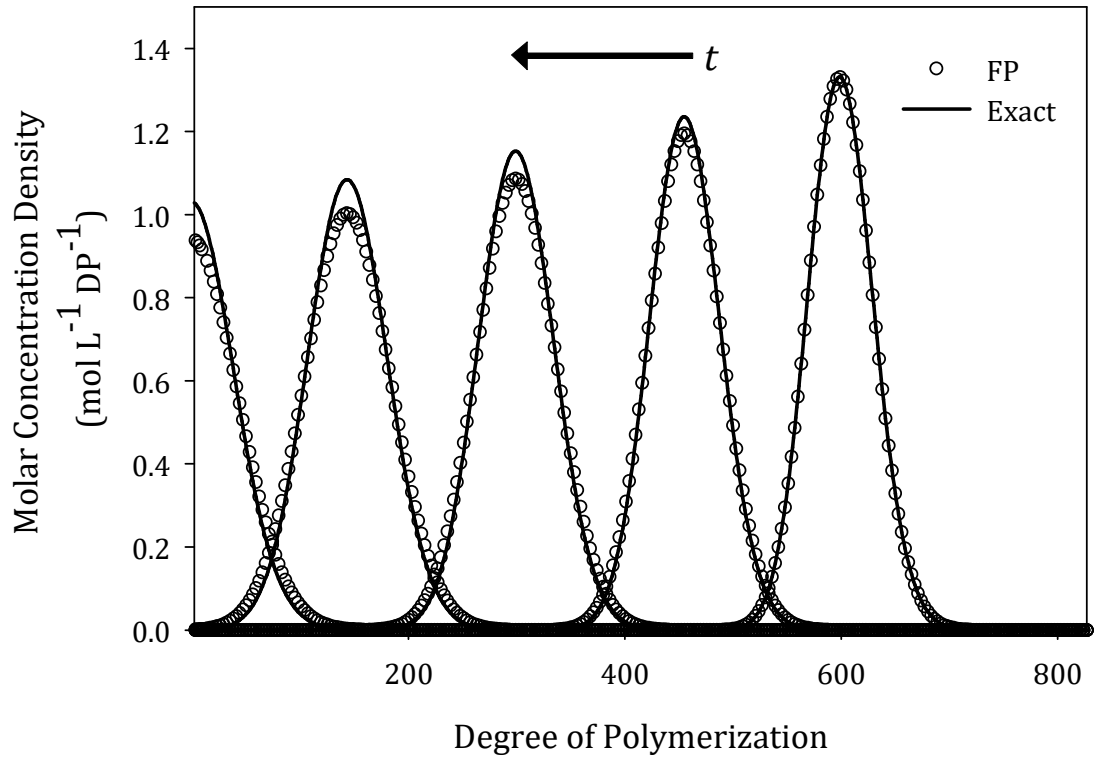


Figure 4.11: Temporal evolution of the molar concentration density using the fixed pivot (FP) and the exact solutions for chain-end scission. The highest peak corresponds to $t = 0$ time units, followed by $t = 144, 300, 456,$ and 600 time units in the order of decreasing peak heights. Here, $[p, q] = [10, 545]$ where p and q are the number of pivots for the discrete and continuous region respectively.

Such observations can be explained using the illustrations in Figure 4.12. For simplicity, $k^\gamma(v)=1$ was employed following Stickel and Griggs (2012). For a non-constant $k^\gamma(v)$ (not shown), the same trend was also observed where over-predictions occurred near the base of the peak and under-predictions occurred near the peak. Figure 4.12a represents the left part of the peak where the slope of the molar concentration density is increasing ($\frac{d^2c}{dx^2} > 0$). Let the quantity $\frac{c_{i+1}-c_i}{x_{i+1}-x_i} = c_{i+1}-c_i$ (with $x_{i+1}-x_i=1$) represent the exact solution for $\frac{dc_i}{dt}$ at a particular instance. With a coarser mesh, a finite difference formula spanning a larger range in x will be used to approximate $\frac{dc_i}{dt}$ at x_i . Had the molar concentration density increased linearly along the v axis, this exact solution would be retained at a coarser mesh as $\frac{c_{i+1}^*-c_i}{x_{i+1}^*-x_i} = \frac{c_{i+1}-c_i}{x_{i+1}-x_i}$. However, because the molar concentration density increases instead, i.e. at x_{i+1}^* , the corresponding molar concentration density is \hat{c}_{i+1} , having $\frac{\hat{c}_{i+1}-c_i}{x_{i+1}^*-x_i} > \frac{c_{i+1}^*-c_i}{x_{i+1}^*-x_i}$ resulted in an over-prediction of $\frac{dc_i}{dt}$ at x_i over time. The opposite happens after the inflection point towards the peak, where the molar concentration density tapers off, i.e. $\frac{d^2c}{dx^2} < 0$ (Figure 4.12b). In this case, $\frac{\hat{c}_{i+1}-c_i}{x_{i+1}^*-x_i} < \frac{c_{i+1}^*-c_i}{x_{i+1}^*-x_i}$, resulting in under-predictions. Similar interpretation extends to the right of the peak: before the inflection point in Figure 4.12c, $\left| \frac{\hat{c}_{i+1}-c_i}{x_{i+1}^*-x_i} \right| > \left| \frac{c_{i+1}^*-c_i}{x_{i+1}^*-x_i} \right|$ resulting in a more negative rate than the exact solution whereas after the inflection point in Figure 4.12d, $\left| \frac{\hat{c}_{i+1}-c_i}{x_{i+1}^*-x_i} \right| < \left| \frac{c_{i+1}^*-c_i}{x_{i+1}^*-x_i} \right|$, resulting in a less negative rate compared to the

exact solution. In short, the FP solution will over-predict c_i when $\frac{d^2c}{dx^2} > 0$ and under-predict otherwise. The discrepancies are magnified chiefly by steep gradients. Mesh refinement improves the solution but then not much reduction in the number of equations can be gained. A more intelligent approach is having selective concentration of the mesh at regions with a steep distribution, as done in the moving pivot method (S. Kumar & Ramkrishna, 1996b) and its extension (Attarakih, Bart, & Faqir, 2003). However, at present evidence for this specific case is not available in the literature. Another alternative is the solution obtained by discretizing and solving the PDE given by Stickel and Griggs (2012) which was shown to attain a good degree of accuracy. Although not explicitly reported in their work, the discretization involving a second order derivative will require care in mesh size, time stepping and the choice of solvers to ensure numerical stability when dealing with such a steep distribution. Despite the inaccuracies in the concentration distributions, the FP technique with an arithmetic mesh ($r = 1$) was able to predict the transients of the zeroth and the first moments of the distribution with $\varepsilon_g < 0.01$ using only 21 pivot points. In some cases moments are of greater interest than the complete distribution (Madras, Smith, et al., 1997; Oyerokun & Vaia, 2012; Rangarajan et al., 1998; Striegel, 2003). Thus, depending on the information required, users encountering a narrow initial polymer distribution may choose between these alternatives. Nevertheless, for solving chain-end scission problems involving broad distributions typical of natural polymers, the FP technique appeared to be very efficient.

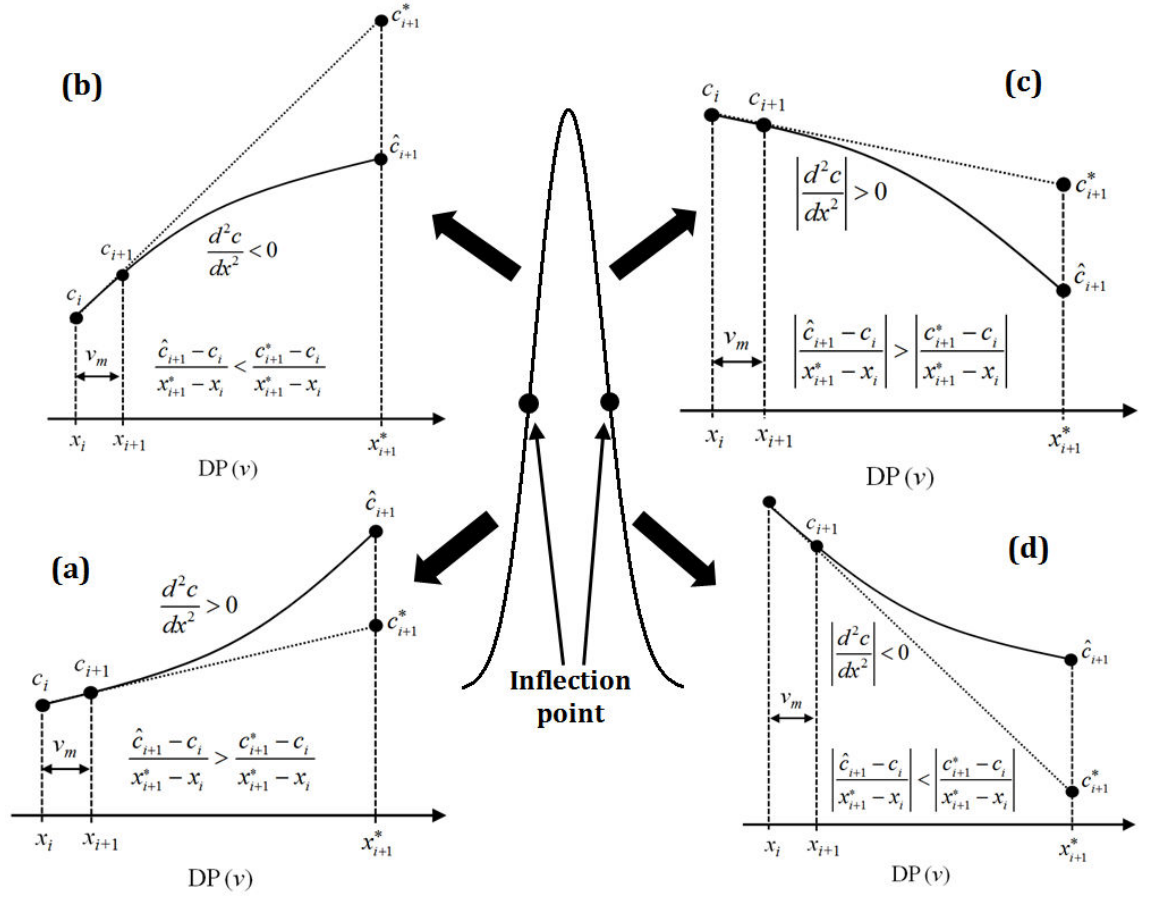


Figure 4.12: Illustrating the limitation of the fixed pivot technique in modelling chain-end scission for steep distributions. Here $v_m = 1$, x_i is the pivot for the i -th interval, and c_i is the molar concentration density. The symbol c_{i+1}^* represents the linearly extrapolated concentration while \hat{c}_{i+1} represents the actual concentration at x_{i+1}^* .

4.1.6 Guidelines for Meshing

Finally, based on the case study presented in Section 4.1.4, the following meshing guidelines are proposed for modelling chain-end scission with the FP technique, assuming that values of N and v_m are typically known:

- (a) Choose $p+q \sim O(10^{-2}N)$. Then choose the value of p such that the constraint of Eq. (4.16) is fulfilled. There is a lower limit for p below which q becomes too large, causing pivots to be spaced closer than v_m . To facilitate this, Figure 4.13 can be

conveniently used to ensure that $p \geq p_{min}$. For more advanced users, the MATLAB script file for calculating p_{min} is given in Section A.6 of Appendix A.

- (b) For values of $p+q$ and N not covered in Figure 4.13, first choose $p = 50 \sim 100$ and calculate q using Eq. (4.16). If the resulting $p+q \sim O(10^{-2}N)$, proceed to solve the system of differential equations. Otherwise, select a different value of p and repeat the procedure until the desired $p+q$ is found.

The meshing guidelines above are to ensure that the continuous region is sufficiently resolved, while at the same time the discrete region is sufficiently wide to meet the requirements of the application, e.g. to enable comparison with experimental measurement of oligomer concentrations. For chain-end scission, typically resolving the continuous region is the deciding factor. The resolution of the discrete region becomes important only when random scission is involved, as will be shown later in Section 4.3.

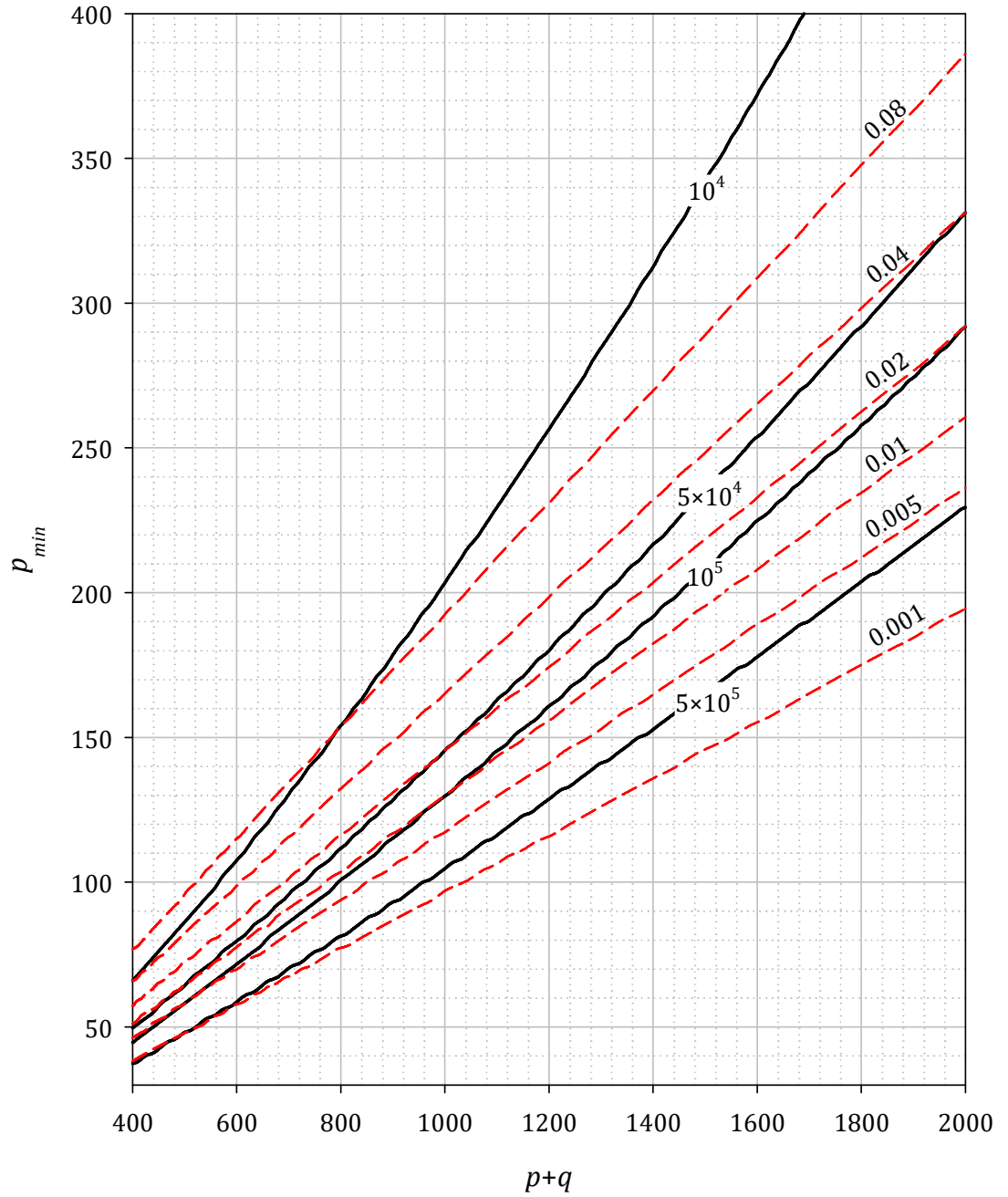


Figure 4.13: Minimum value of p (p_{min}) at a selected value of $p+q$ for a given N (solid line), with $v_m = 1$. The dashed lines correspond to $(p+q)/N$.

4.2 Modelling Simultaneous Random and Chain-End Scissions

As alluded to in the literature survey (Section 2.2.1), the FP solution for PBEs with the uniform random scission kernel using a discrete-continuous split of the DP domain had been attempted (S. Kumar & Ramkrishna, 1996a). However, as the FP framework for

chain-end scission had not been fully established prior to this, modelling of simultaneous random and chain-end scissions within a unified FP framework had not been attempted. Establishment of such a framework is important in working towards the ultimate goal of predicting simultaneous random and chain-end scissions for enzymatic depolymerization.

4.2.1 Fully Discrete (Exact) Solution for Simultaneous Random and Chain-End Scissions

Referring to Eqs. (2.4) - (2.5) of Section 2.2.1, the fully discrete PBE for random scission with a uniform scission kernel and first order kinetics can be compactly rewritten in the following form:

Pure random scission:

$$\frac{dC_1(t)}{dt} = \sum_{j=2}^N k_j^\alpha \left(\frac{2}{j-1} \right) C_j(t) \quad i = 2, 3, \dots, N \quad (4.38)$$

$$\frac{dC_i(t)}{dt} = \sum_{j=i+1}^N k_j^\alpha \left(\frac{2}{j-1} \right) C_j(t) - k_i^\alpha C_i(t) \quad i = 2, 3, \dots, N-1 \quad (4.39)$$

$$\frac{dC_N(t)}{dt} = -k_N^\alpha C_N(t) \quad (4.40)$$

For ease of reading, the equations for chain-end scission are reproduced below:

Pure chain-end scission:

$$\frac{dC_1(t)}{dt} = 2k_2^\gamma C_2(t) + \sum_{j=3}^N k_j^\gamma C_j(t) \quad (4.41)$$

$$\frac{dC_i(t)}{dt} = k_{i+1}^\gamma C_{i+1}(t) - k_i^\gamma C_i(t) \quad i = 2, 3, \dots, N-1 \quad (4.42)$$

$$\frac{dC_N(t)}{dt} = -k_N^\gamma C_N(t) \quad (4.43)$$

As opposed to the equations for chain-end scission, where only polymers with $DP = i+1$ contribute to the birth term of $DP = i$, for random scission the birth term of $DP = i$ is essentially a recipient of all breakages occurring from $DP = i+1$ and beyond.

Having now established the fully discrete solution for both chain-end and random scissions, the fully discrete solution for simultaneously occurring random and chain-end scissions is given simply as:

Simultaneous random and chain-end scissions:

$$\frac{dC_1(t)}{dt} = 2k_2^\gamma C_2(t) + \sum_{j=3}^N k_j^\gamma C_j(t) + \sum_{j=2}^N k_j^\alpha \left(\frac{2}{j-1} \right) C_j(t) \quad i = 2, 3, \dots, N \quad (4.44)$$

$$\frac{dC_i(t)}{dt} = k_{i+1}^\gamma C_{i+1}(t) + \sum_{j=i+1}^N k_j^\alpha \left(\frac{2}{j-1} \right) C_j(t) - k_i^\gamma C_i(t) - k_i^\alpha C_i(t) \quad i = 2, 3, \dots, N-1 \quad (4.45)$$

$$\frac{dC_N(t)}{dt} = -k_N^\gamma C_N(t) - k_N^\alpha C_N(t) \quad (4.46)$$

The integration of Eqs. (4.44) - (4.46) by using commercial ODE solvers intended for general usage is inefficient at large values of N . However, as these equations are the golden standard for solving simultaneous random and chain-end scissions, the ability to solve them within a reasonable amount of time is important. This is dealt with in the next section.

4.2.2 Computing the Exact Solution for the Purpose of Validation

In general, Eqs. (4.44) - (4.46) form a linear system of ODEs which can be written in the following form:

$$\frac{d\mathbf{C}(t)}{dt} = \mathbf{Z}\mathbf{C}(t) \quad (4.47)$$

where $\mathbf{C}^T = [C_1 \ C_2 \ C_3 \ \cdots \ C_N]$ and \mathbf{Z} is the upper triangular coefficient matrix given by:

$$\mathbf{Z} = \begin{bmatrix} 0 & 2k_2^\gamma + 2k_2^\alpha & k_3^\gamma + k_3^\alpha & k_4^\gamma + \frac{2}{3}k_4^\alpha & k_5^\gamma + \frac{2}{4}k_5^\alpha & \cdots & k_{N-1}^\gamma + \frac{2}{N-2}k_{N-1}^\alpha & k_N^\gamma + \frac{2}{N-1}k_N^\alpha \\ 0 & -(k_2^\gamma + k_2^\alpha) & k_3^\gamma + k_3^\alpha & \frac{2}{3}k_4^\alpha & \frac{2}{4}k_5^\alpha & \cdots & \frac{2}{N-2}k_{N-1}^\alpha & \frac{2}{N-1}k_N^\alpha \\ \vdots & & -(k_3^\gamma + k_3^\alpha) & k_4^\gamma + \frac{2}{3}k_4^\alpha & \frac{2}{4}k_5^\alpha & \cdots & \vdots & \vdots \\ \vdots & & & -(k_4^\gamma + k_4^\alpha) & k_5^\gamma + \frac{2}{4}k_5^\alpha & \ddots & \vdots & \vdots \\ \vdots & & & & -(k_5^\gamma + k_5^\alpha) & \ddots & \frac{2}{N-2}k_{N-1}^\alpha & \vdots \\ \vdots & & \mathbf{0} & & & \ddots & k_{N-1}^\gamma + \frac{2}{N-2}k_{N-1}^\alpha & \frac{2}{N-1}k_N^\alpha \\ \vdots & & & & & & -(k_{N-1}^\gamma + k_{N-1}^\alpha) & k_N^\gamma + \frac{2}{N-1}k_N^\alpha \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & -(k_N^\gamma + k_N^\alpha) \end{bmatrix} \quad (4.48)$$

Compared to the case of pure chain-end scission, i.e. by removing all the terms with the superscript ‘ α ’ in Eq. (4.48), the coefficient matrix \mathbf{Z} involving random scission contains more non-zero entries. Particularly when the value of N is large, e.g. $\sim O(10^4)$ and beyond, the integration of Eqs. (4.44) - (4.46) by commercial integrators using the default settings is inefficient. Although commercial implicit solvers (e.g. the ‘ode15s’ subroutine in MATLAB[®] 2009b) may have the option of receiving the Jacobian matrix (which equals \mathbf{Z} in this case) as an input to speed up computation, in general a system involving random scission will require more computational time than a corresponding system for chain-end scission because more equations are solved by the solver at each integration step. In addition, as PBEs with a random scission kernel may exhibit solutions with distinct time scales (e.g. C_N may increase and disappear rapidly at the onset of simulation but the accumulation of C_5 may only be significant towards the end of the simulation etc.) resulting in a very stiff system of ODEs, the use of explicit solvers (e.g. the ‘ode45’ subroutine in MATLAB[®] 2009b) would be inefficient. Because the validation

of the FP solution (presented later in Section 4.2.3) with the exact solution is a critical step to enable confident usage of the FP technique, here the linear structure of Eqs. (4.44) - (4.46) and the fact that \mathbf{Z} is upper triangular are exploited to enable the system of ODEs to be solved more efficiently using a generic commercial ODE solver on a typical workstation.

To begin, the basic idea here is to break the system of ODEs into smaller pieces which are more tractable by the ODE solver. This can be done by introducing partitions into the system of ODEs as shown here:

$$\frac{d}{dt} \begin{bmatrix} \mathbf{C}^{(1)}(t) \\ \mathbf{C}^{(2)}(t) \\ \vdots \\ \mathbf{C}^{(w-1)}(t) \\ \mathbf{C}^{(w)}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{G}^{(1,1)} & \mathbf{F}^{(1,2)} & \dots & \dots & \mathbf{F}^{(1,w-1)} & \mathbf{F}^{(1,w)} \\ & \mathbf{G}^{(2,2)} & \mathbf{F}^{(2,3)} & \dots & \mathbf{F}^{(2,w-1)} & \mathbf{F}^{(2,w)} \\ & & \ddots & & \vdots & \vdots \\ & & & \mathbf{G}^{(w-2,w-2)} & \mathbf{F}^{(w-2,w-1)} & \mathbf{F}^{(w-2,w)} \\ & \mathbf{0} & & & \mathbf{G}^{(w-1,w-1)} & \mathbf{F}^{(w-1,w)} \\ & & & & & \mathbf{G}^{(w,w)} \end{bmatrix} \begin{bmatrix} \mathbf{C}^{(1)}(t) \\ \mathbf{C}^{(2)}(t) \\ \vdots \\ \mathbf{C}^{(w-1)}(t) \\ \mathbf{C}^{(w)}(t) \end{bmatrix} \quad (4.49)$$

where $[\mathbf{C}^{(i)}]^T = [C_{S(i-1)+1} \ C_{S(i-1)+2} \ \dots \ C_{S(i-1)+S}]$ and $i = 1, 2, \dots, w$. Here, w is an integer which represents the total number of partitions for the molar concentration vector which can be obtained as $w = N/S$ with S being the size of the individual partition. The constant square matrices $\mathbf{G}^{(i,i)}$ as well as $\mathbf{F}^{(i,j)}$ ($i = 1, 2, \dots, w, j = i+1, i+2, \dots, w$) are given as:

$$\mathbf{G}^{(i,i)} = \begin{bmatrix} Z_{S(i-1)+1, S(i-1)+1} & Z_{S(i-1)+1, S(i-1)+2} & Z_{S(i-1)+1, S(i-1)+3} & \dots & Z_{S(i-1)+1, S(i-1)+S-1} & Z_{S(i-1)+1, S(i-1)+S} \\ & Z_{S(i-1)+2, S(i-1)+2} & Z_{S(i-1)+2, S(i-1)+3} & \dots & Z_{S(i-1)+2, S(i-1)+S-1} & Z_{S(i-1)+2, S(i-1)+S} \\ & & Z_{S(i-1)+3, S(i-1)+3} & \dots & Z_{S(i-1)+3, S(i-1)+S-1} & Z_{S(i-1)+3, S(i-1)+S} \\ & & & \ddots & \vdots & \vdots \\ & \mathbf{0} & & & Z_{S(i-1)+S-1, S(i-1)+S-1} & \vdots \\ & & & & & Z_{S(i-1)+S, S(i-1)+S} \end{bmatrix} \quad (4.50)$$

$$\mathbf{F}^{(i,j)} = \begin{bmatrix} Z_{S(i-1)+1,S(j-1)+1} & Z_{S(i-1)+1,S(j-1)+2} & \cdots & Z_{S(i-1)+1,S(j-1)+S} \\ Z_{S(i-1)+2,S(j-1)+1} & Z_{S(i-1)+2,S(j-1)+2} & \cdots & Z_{S(i-1)+2,S(j-1)+S} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{S(i-1)+S,S(j-1)+1} & Z_{S(i-1)+S,S(j-1)+2} & \cdots & Z_{S(i-1)+S,S(j-1)+S} \end{bmatrix} \quad (4.51)$$

As the square matrices $\mathbf{G}^{(i,i)}$ and $\mathbf{F}^{(i,j)}$ are stationary in time, it follows that the molar concentration vector can be computed iteratively using backward substitution. Beginning at $i = w$, the following ODE can be readily integrated by numerical solvers:

$$\frac{d}{dt} \mathbf{C}^{(w)}(t) = \mathbf{G}^{(w,w)} \mathbf{C}^{(w)}(t) \quad (4.52)$$

Because of the nature of numerical ODE solvers, solution of Eq. (4.52) yields discrete data points. For the sake of discussion, the resulting solution is referred to as $\mathbf{C}^{(w)}(t_k)$ ($k = 0, 1, 2, \dots$) where t_k is the time at which the solution is evaluated and the step size taken by the solver is $\Delta t = t_{k+1} - t_k$. Having obtained $\mathbf{C}^{(w)}(t_k)$, it is retained for the next calculation loop to compute $\mathbf{C}^{(w-1)}(t_k)$. At $i = w - 1$:

$$\frac{d}{dt} \mathbf{C}^{(w-1)}(t) = \mathbf{G}^{(w-1,w-1)} \mathbf{C}^{(w-1)}(t) + \mathbf{F}^{(w-1,w)} \mathbf{C}^{(w)}(t) \quad (4.53)$$

Now that $\mathbf{C}^{(w)}(t_k)$ is readily available, when dealing with Eq. (4.53), the ODE solver need not solve for $\mathbf{C}^{(w)}(t_k)$ again, but that the second term on the RHS, i.e. $\mathbf{F}^{(w-1,w)} \mathbf{C}^{(w)}(t)$, can be approximated by linearly interpolating the $\mathbf{C}^{(w)}(t_k)$ retained from the previous calculation loop. Similar procedure is repeated for the computation of $\mathbf{C}^{(w-2)}(t_k), \mathbf{C}^{(w-3)}(t_k), \dots, \mathbf{C}^{(1)}(t_k)$. In short, the following ODE is numerically solved in an iterative fashion for $i = w$ to 1:

$$\frac{d}{dt} \mathbf{C}^{(i)}(t) = \mathbf{G}^{(i,i)} \mathbf{C}^{(i)}(t) + \sum_{j=i+1}^w \mathbf{F}^{(i,j)} \mathbf{C}^{(j)}(t) \quad (4.54)$$

where $\mathbf{C}^{(j)}(t)$ is approximated by interpolating the $\mathbf{C}^{(j)}(t_k)$ brought forward from the previous computation loops. In contrast to solving the full sized system where the ODE solver is forced to take the same number of integration steps for $DP = 1$ to $DP = N$, the procedure described here allows every smaller problem to be integrated with different number of time steps. As the stiffness of the smaller system of ODEs is much lesser than that of the full sized system of ODEs (due to the grouping of polymers which evolve at similar time scales), allowing each smaller system to be solved at different number of time steps is much more efficient. To summarize, the procedure is given graphically in Figure 4.14.

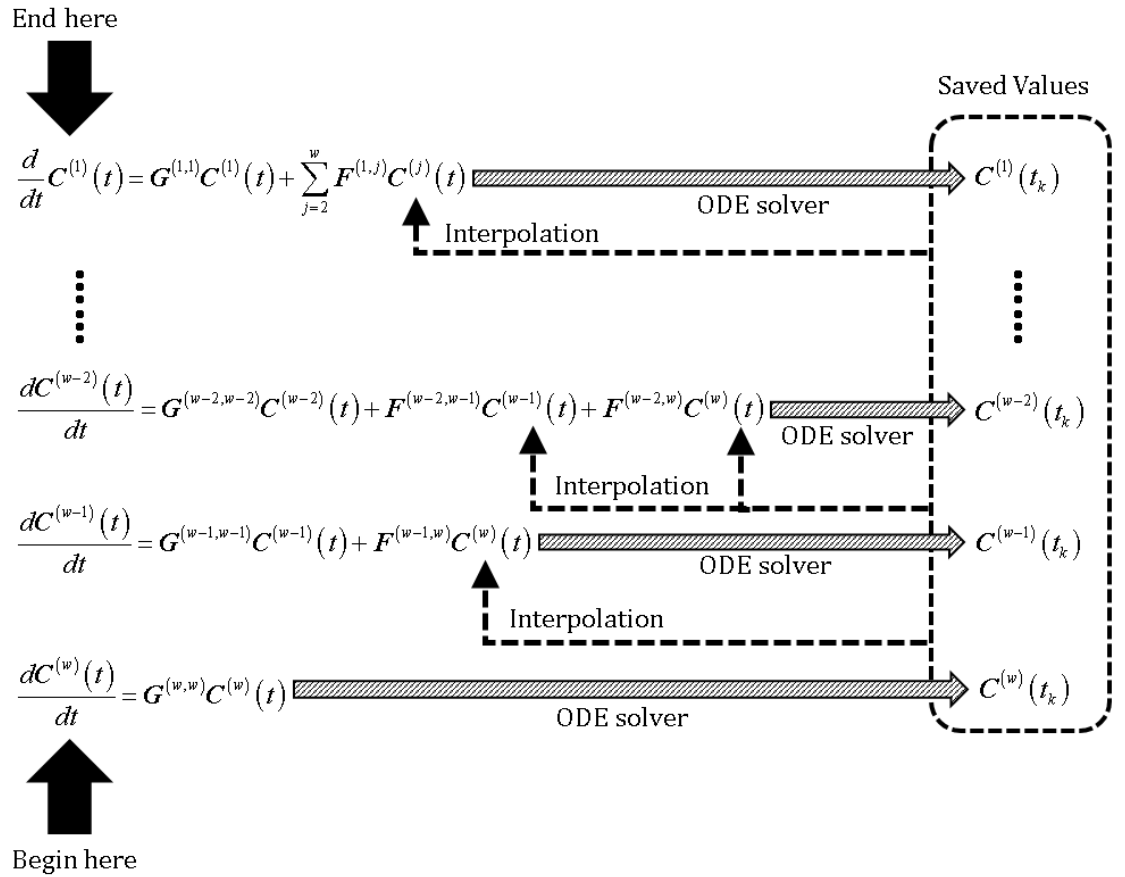


Figure 4.14: Illustrating the iterative procedure for obtaining $\mathbf{C}^{(i)}(t_k)$, $i = 1$ to w .

A numerical demonstration of the reasons for the efficacy of the proposed method will be given in a case study in Section 4.2.4.

4.2.3 Fixed Pivot Discretization for Simultaneous Random and Chain-End Scissions

To allow the prediction of simultaneously occurring random and chain-end scissions using the FP technique, the key is to superimpose the equations for random and chain-end scissions on the same mesh, i.e. by partitioning the DP axis into p pivot points in the discrete region and q pivot points in the continuous region with their values being constrained by Eq. (4.16). As derived in detail in Section B.1 of Appendix B, for pure random scission [Eqs. (4.38) - (4.40)], the FP approximation for the fully discrete solution with such a mesh is as follows:

Pure random Scission:

Discrete region:

$$\frac{dC_1(t)}{dt} = \sum_{j=2}^{p+q} k_j^\alpha \left[\frac{2}{x_j - 1} \right] C_j(t) \quad (4.55)$$

$$\frac{dC_i(t)}{dt} = \sum_{j=i+1}^{p+q} k_j^\alpha \left[\frac{2}{x_j - 1} \right] C_j(t) - k_i^\alpha C_i(t) \quad i = 2, 3, \dots, p-1 \quad (4.56)$$

Discrete-continuous boundary:

$$\frac{dC_p(t)}{dt} = \sum_{j=p+1}^{p+q} k_j^\alpha \left[\frac{2}{x_j - 1} \right] C_j(t) - k_p^\alpha C_p(t) \quad (4.57)$$

$$\frac{dC_{p+1}(t)}{dt} = \sum_{j=p+2}^{p+q} k_j^\alpha \left[\frac{x_{p+2} - x_{p+1}}{x_j - 1} \right] C_j(t) - k_{p+1}^\alpha C_{p+1}(t) \quad (4.58)$$

Continuous region:

$$\frac{dC_i(t)}{dt} = \left[\frac{x_i - x_{i-1}}{x_i - 1} \right] k_i^\alpha C_i(t) + \sum_{j=i+1}^{p+q} \left[\frac{x_{i+1} - x_i}{x_j - 1} + \frac{x_i - x_{i-1}}{x_j - 1} \right] k_j^\alpha C_j(t) - k_i^\alpha C_i(t) \quad (4.59)$$

$i = p+2, p+3, \dots, p+q-1$

$$\frac{dC_{p+q}(t)}{dt} = \left[\frac{x_{p+q} - x_{p+q-1}}{x_{p+q} - 1} \right] k_{p+q}^\alpha C_{p+q}(t) - k_{p+q}^\alpha C_{p+q}(t) \quad (4.60)$$

Having established the above, it follows that the FP approximation for the simultaneous random and chain-end scissions is the result of superimposing Eqs. (4.23) - (4.28) (for chain-end scission) and Eqs. (4.55) - (4.60) (for random scission), as shown below:

Simultaneous random and chain-end scissions:

Discrete region:

$$\frac{dC_1(t)}{dt} = 2k_2^\gamma C_2(t) + \sum_{j=3}^{p+q} k_j^\gamma C_j(t) + \sum_{j=2}^{p+q} k_j^\alpha \left[\frac{2}{x_j - 1} \right] C_j(t) \quad (4.61)$$

$$\frac{dC_i(t)}{dt} = k_{i+1}^\gamma C_{i+1}(t) + \sum_{j=i+1}^{p+q} k_j^\alpha \left[\frac{2}{x_j - 1} \right] C_j(t) - k_i^\gamma C_i(t) - k_i^\alpha C_i(t) \quad i = 2, 3, \dots, p-1 \quad (4.62)$$

Discrete-continuous boundary:

$$\frac{dC_p(t)}{dt} = k_{p+1}^\gamma C_{p+1}(t) + \sum_{j=p+1}^{p+q} k_j^\alpha \left[\frac{2}{x_j - 1} \right] C_j(t) - k_p^\gamma C_p(t) - k_p^\alpha C_p(t) \quad (4.63)$$

$$\frac{dC_{p+1}(t)}{dt} = \left[\frac{v_m}{x_{p+2} - x_{p+1}} \right] k_{p+2}^\gamma C_{p+2}(t) + \sum_{j=p+2}^{p+q} k_j^\alpha \left[\frac{x_{p+2} - x_{p+1}}{x_j - 1} \right] C_j(t) - k_{p+1}^\gamma C_{p+1}(t) - k_{p+1}^\alpha C_{p+1}(t) \quad (4.64)$$

Continuous region:

$$\begin{aligned} \frac{dC_i(t)}{dt} = & \left[\frac{v_m}{x_{i+1} - x_i} \right] k_{i+1}^\gamma C_{i+1}(t) + \left[1 - \frac{v_m}{x_i - x_{i-1}} \right] k_i^\gamma C_i(t) + \left[\frac{x_i - x_{i-1}}{x_i - 1} \right] k_i^\alpha C_i(t) \\ & + \sum_{j=i+1}^{p+q} \left[\frac{x_{i+1} - x_i}{x_j - 1} + \frac{x_i - x_{i-1}}{x_j - 1} \right] k_j^\alpha C_j(t) - k_i^\gamma C_i(t) - k_i^\alpha C_i(t) \end{aligned} \quad (4.65)$$

$$i = p+2, p+3, \dots, p+q-1$$

$$\begin{aligned} \frac{dC_{p+q}(t)}{dt} = & \left[1 - \frac{v_m}{x_{p+q} - x_{p+q-1}} \right] k_{p+q}^\gamma C_{p+q}(t) + \left[\frac{x_{p+q} - x_{p+q-1}}{x_{p+q} - 1} \right] k_{p+q}^\alpha C_{p+q}(t) \\ & - k_{p+q}^\gamma C_{p+q}(t) - k_{p+q}^\alpha C_{p+q}(t) \end{aligned} \quad (4.66)$$

As ultimately it is the desired outcome of this study to predict the enzymatic depolymerization phenomena involving random and chain-end scissions, the set of Eqs. (4.61) - (4.66) is the critical workhorse for the efficient solution of PBEs. Section 4.2.4 will be devoted to validating the performance of the FP solution against the exact solutions given by Eqs. (4.44) – (4.46) in Section 4.2.1.

4.2.4 Case Study on Simultaneously Occurring Random and Chain-End Scissions

For the purpose of validation, the polymer system described in Section 4.1.4 for chain-end scission was used here. Recalling briefly, the initial condition is specified by Eqs. (3.1) - (3.3), where $N = 22496$, $m_S(0) = 10$ g/L, $\overline{M}_n = 4100$, and $\overline{M}_w = 5430$. Here, the DP-dependent rate kernel (S. Kumar & Ramkrishna, 1996a; Staggs, 2004) for chain-end scission used previously (i.e. $k^\gamma(v) = k_p^\gamma v$) was retained. For random scission, the appropriate power law rate kernel is of the form of $k^\alpha(v) = k_p^\alpha (v-1)$ according to S. Kumar and Ramkrishna (1996a). As practiced previously, the values of k_p^γ and k_p^α were chosen arbitrarily as 1/80 and 1/200 respectively such that a total evolution time of around 600 was achieved. Since the ultimate goal here is to validate the efficacy of the FP technique in solving simultaneous random and chain-end scissions, the optimal time

duration for solving the fully discrete PBEs via the procedure presented in Section 4.2.2 is relatively unimportant as long as the exact solution can be obtained in a reasonable amount of time. For the case study described here, since the exact solution can be obtained in less than 10 minutes by dividing the molar concentration vector into $w = 19$ partitions with each partition having a size of $S = 1184$, these values were used in all the following simulations. From the previous results presented for chain-end scission, values of $[p, q] = [100, 500]$ which totalled less than 3% of the N equations used in the exact solution were found to be adequate, and thus retained here. Because it was found during preliminary simulations that the use of ‘ode45’ sub-routine was relatively inefficient, the ‘ode15s’ sub-routine of MATLAB[®] 2009b was used here and all computations in this case study were done on workstation C. Example codes for the simulation of simultaneous random and chain-end scissions using the FP and the exact solution are given in Appendix C.

Before examining the FP solution with the exact solution, the cause of the improved efficiency in solving the fully discrete PBEs is shown briefly. A useful indicator for the stiffness of a linear ODE system is the stiffness ratio (Aiken, 1985), defined as:

$$\text{Stiffness ratio} = \frac{\max(|\lambda_i|)}{\min(|\lambda_i|)} \quad (4.67)$$

where λ_i is the vector of eigenvalues for $\mathbf{G}^{(i,i)}$. The minimum value of the stiffness ratio is one and the stiffness of a system of ODEs increases with the increasing magnitude of the stiffness ratio. As $G_{11}^{(1,1)} = 0$ and consequently an eigenvalue of zero exists, the stiffness ratio for partition 1 was calculated by omitting the zero eigenvalue (otherwise stiffness ratio = ∞). Likewise, for the full sized system, the zero eigenvalue due to $Z_{11} = 0$ was similarly omitted. Figure 4.15 shows the stiffness ratio for the i -th ODE system ($i = 1$ to w) for a total number of partitions $w = 19$ and $S = 1184$. From the figure, the stiffness

ratio decreases as the value of i increases, i.e. the stiffness of the system of ODEs increases from that for describing $d\mathbf{C}^{(w)}/dt$ to that for $d\mathbf{C}^{(1)}/dt$. By breaking the original system of ODEs into smaller sub-problems, the stiffness ratio for each individual problem ($i \geq 2$) is at least three orders of magnitude smaller than the stiffness ratio for the full sized system of ODEs. Even at $i = 1$, the stiffness ratio is still lesser than a tenth of that for the full sized system. On closer examination, most of the stiffness ratio are between the values of 1 to 2, which implies that the values of the maximum and the minimum eigenvalues are not far apart. In this case where there are no degenerate eigenvalues, i.e. the diagonal elements of $\mathbf{G}^{(i,i)}$ are all dissimilar, the close distance between the maximum and the minimum eigenvalues suggests that each smaller sub-system evolves at a similar time scale. As opposed to solving the full sized system of $N = 22946$ ODEs, which is infeasible even for a single run, the divide-and-conquer method proposed here took only 10 minutes to solve the case study.

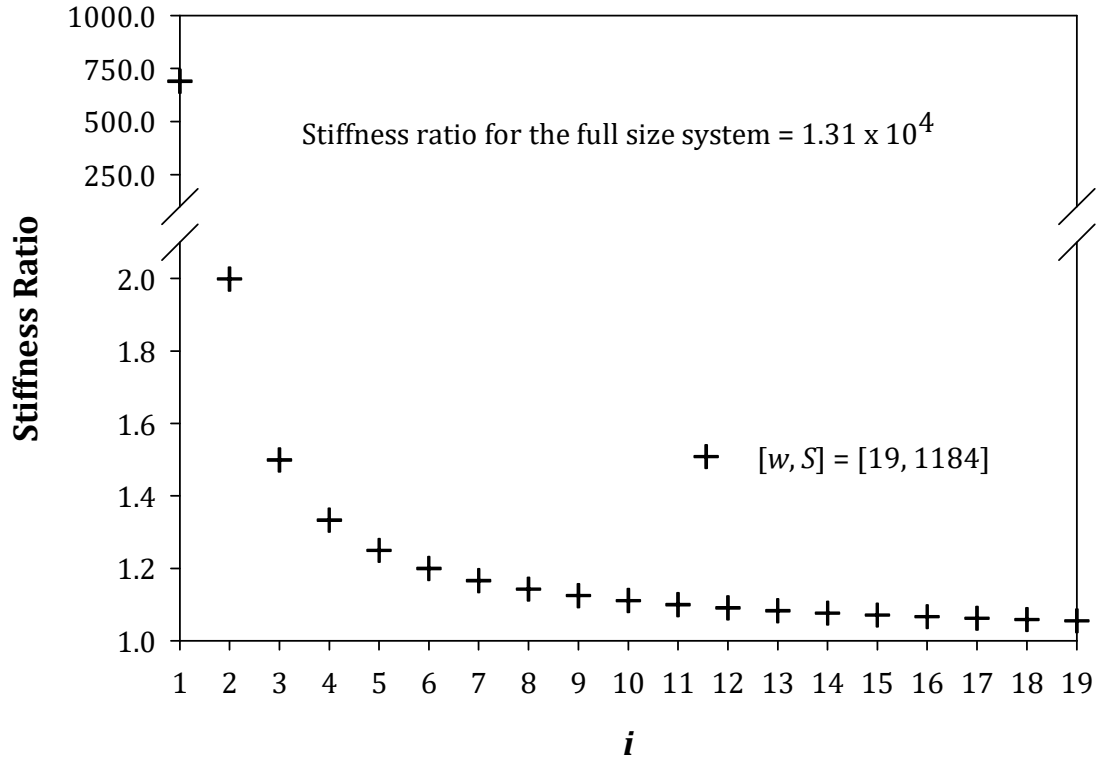


Figure 4.15: Stiffness ratio ($= \max(|\lambda_i|) / \min(|\lambda_i|)$) evaluated from $i = 1$ to w , where λ_i is the vector of eigenvalues for $\mathbf{G}^{(i,i)}$, w is the total number of partitions for the molar concentration vector, and S is the size of each partitioned molar concentration vector.

As was done previously for chain-end scission, it is a good practice to always verify that the conservation of mass is obeyed via the monomer equation. Figure 4.16 shows the transient for the mass concentration of monomer glucose for the case of simultaneous random and chain-end scissions using the exact and the FP solutions. From the figure, a final mass concentration of glucose $= 1.11m_S(0)$ was observed for both solutions, indicating that mass was indeed conserved. Not only that, the FP solution also correctly predicted the transient of the monomer concentration prior to achieving steady state. Here, the error indicator ε_g [Eq. (3.8)] is $\sim O(10^{-3})$. Although the error accrued here is larger than that for pure chain-end scission, i.e. $\varepsilon_g \sim O(10^{-6})$ described previously, it is nonetheless still negligible compared to the order of magnitude of the monomer concentration.

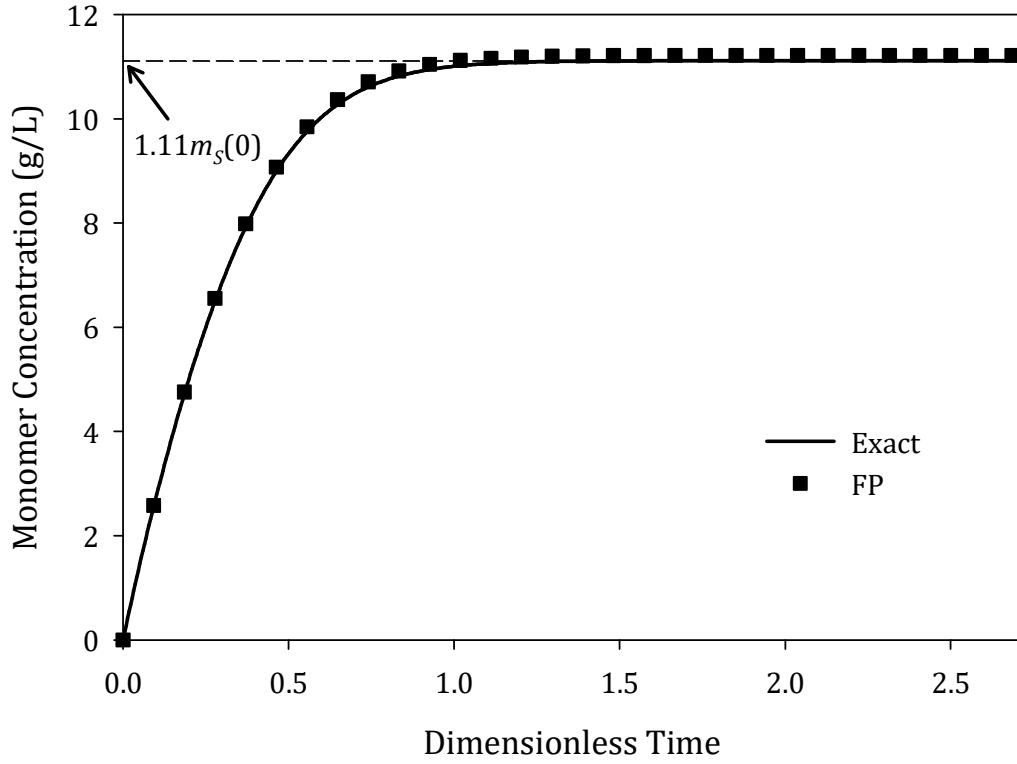


Figure 4.16: Transient of the mass concentration of monomer (glucose) using the exact and fixed pivot (FP) solutions for simultaneous random and chain-end scissions. Here $[p, q] = [100, 500]$, $r = 1.0109$ and $m_s(0) = 10$ g/L. The dimensionless time is normalized against the time required for 99% monomer production and the error in the initial mass due to discretization $[\varepsilon_s, \text{cf. Eq. (3.9)}]$ is $2.61 \times 10^{-4} m_s(0)$.

The concentration profiles of the corresponding DP2 – DP7 oligomers are shown in Figure 4.17. The results show that the transients of the mass concentration of oligomers for simultaneous random and chain-end scissions were predicted well with reference to the exact solutions, with $\varepsilon_g = 1.5 \times 10^{-4}$ and 3.9×10^{-5} for DP2 and DP7 respectively. As opposed to pure chain-end scission where the concentrations of these oligomers do not build up significantly until the late phase of the reaction, in this case the co-existence of random scission resulted in the accumulation of these oligomers even at the onset of simulation. This relatively larger built-up of DP2 – DP7 oligomers, i.e. $\sim O(10^{-2})$ to

$\sim O(10^{-1})$, is much larger than ε_g and therefore ε_g of $\sim O(10^{-5})$ to $\sim O(10^{-4})$ is relatively insignificant as far as the accuracy of the solution is concerned.

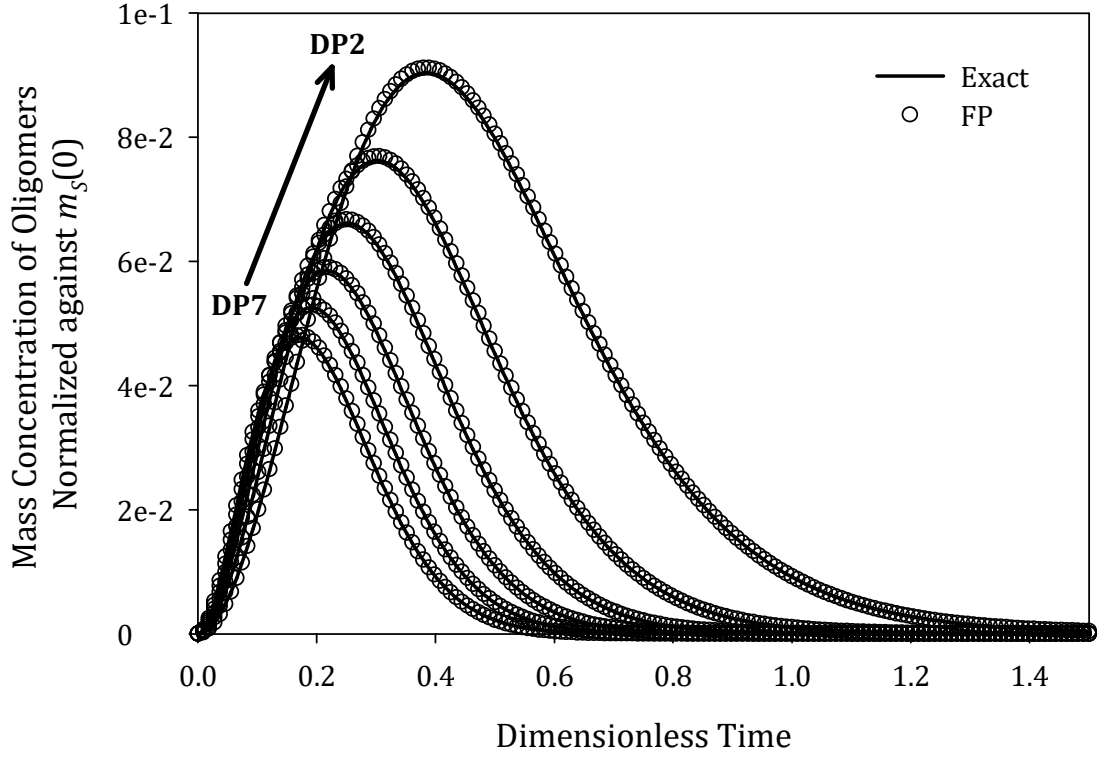


Figure 4.17: Transients of the normalized mass concentration of oligomers using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. The highest peak corresponds to the DP2 oligomer, followed by the DP3 – DP7 oligomers in the order of decreasing peak heights. Here $[p, q] = [100, 500]$, $r = 1.0109$, $m_s(0) = 10$ g/L, and the error in the initial mass due to discretization (ε_s) is $2.61 \times 10^{-4} m_s(0)$. The dimensionless time is normalized against the time required for 99% monomer production.

Figure 4.18 further shows the temporal evolution of the molar concentration density at different dimensionless times. Again, accurate predictions were obtained by the FP solution at both the early and the late phase of reactions. Particularly at $\theta > 1$ where more than 99% monomer had been produced, the accuracy attained is remarkable. As mentioned by S. Kumar and Ramkrishna (1996a), the approximation of the fully discrete

solution for random scission is a very challenging problem due to the discontinuous kernel (see Section B.1 of Appendix B) used in formulating the FP solution. This study stretches the limit of the FP technique by successfully simulating simultaneously occurring random and chain-end scissions to a good degree of accuracy.

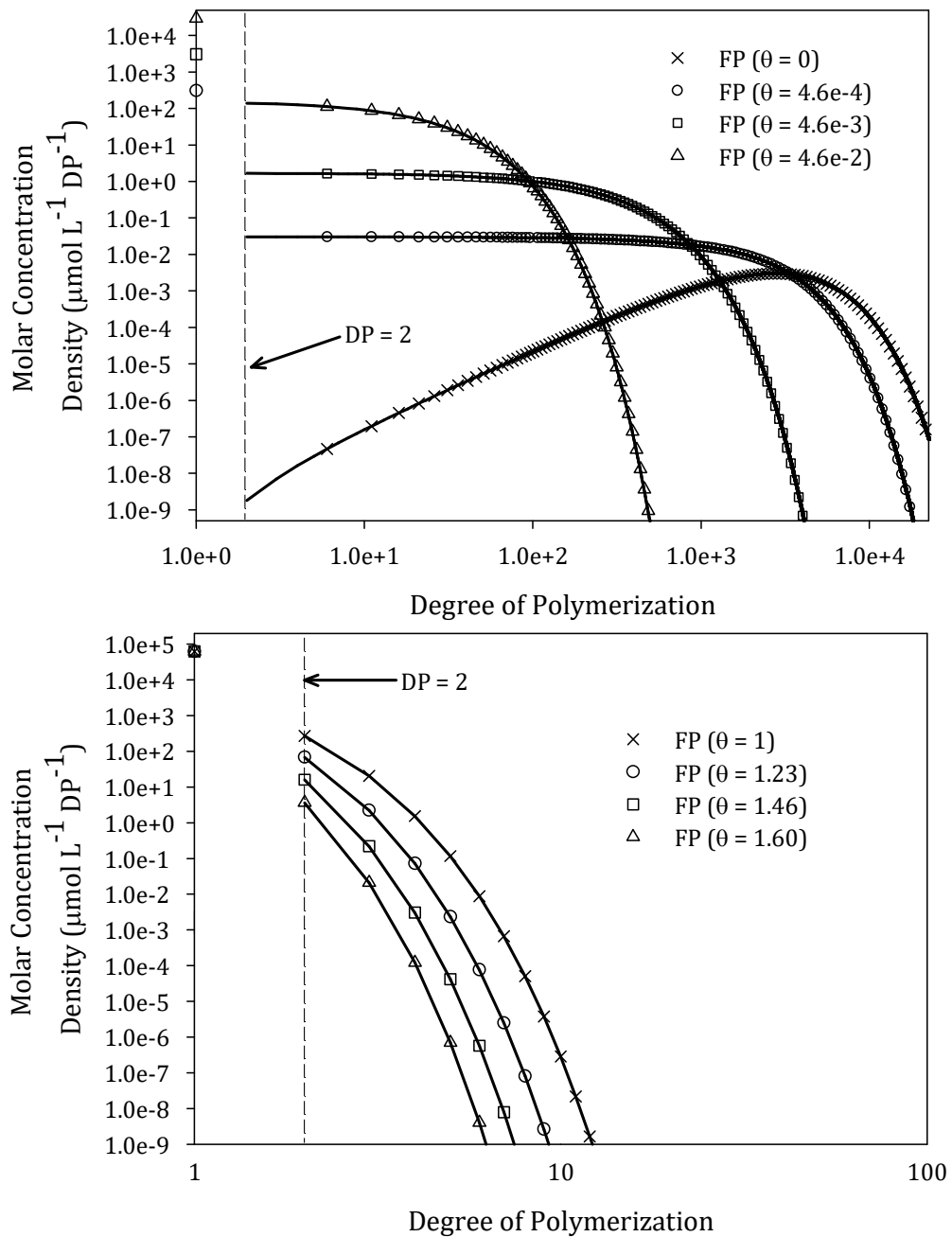


Figure 4.18: Transients of the molar concentration density using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. The dimensionless time (θ) is normalized against the time required for 99% monomer production. Here $[p, q] = [100, 500]$ and $r = 1.0109$. The exact solutions fall on the FP solutions for $\text{DP} = 1$.

Although the results presented thus far are sufficient to demonstrate the efficacy of the FP technique in predicting simultaneous random and chain-end scissions and can be confidently employed for modelling the transient of various polymer species in the SSF

processes (which will be presented in the next chapter), the following results are presented for the sake of completeness. The FP and exact solutions for the temporal evolutions of the number-average DP (\overline{M}_n , Eq. (4.29)), weight-average DP (\overline{M}_w , Eq. (4.30)) and polydispersity index (PD, Eq. (4.31)) for simultaneous random and chain-end scissions are shown in Figures 4.19 – 4.21. From the figures, all these moment-related quantities were predicted accurately, with $\varepsilon_g = 3.3 \times 10^{-4}$ (for \overline{M}_n), $\varepsilon_g = 7.2 \times 10^{-4}$ (for \overline{M}_w), and $\varepsilon_g = 3.1 \times 10^{-4}$ (for PD). As had been elaborated previously for chain-end scission, the results obtained here are not surprising, as the FP technique had been known to be effective in conserving moments.

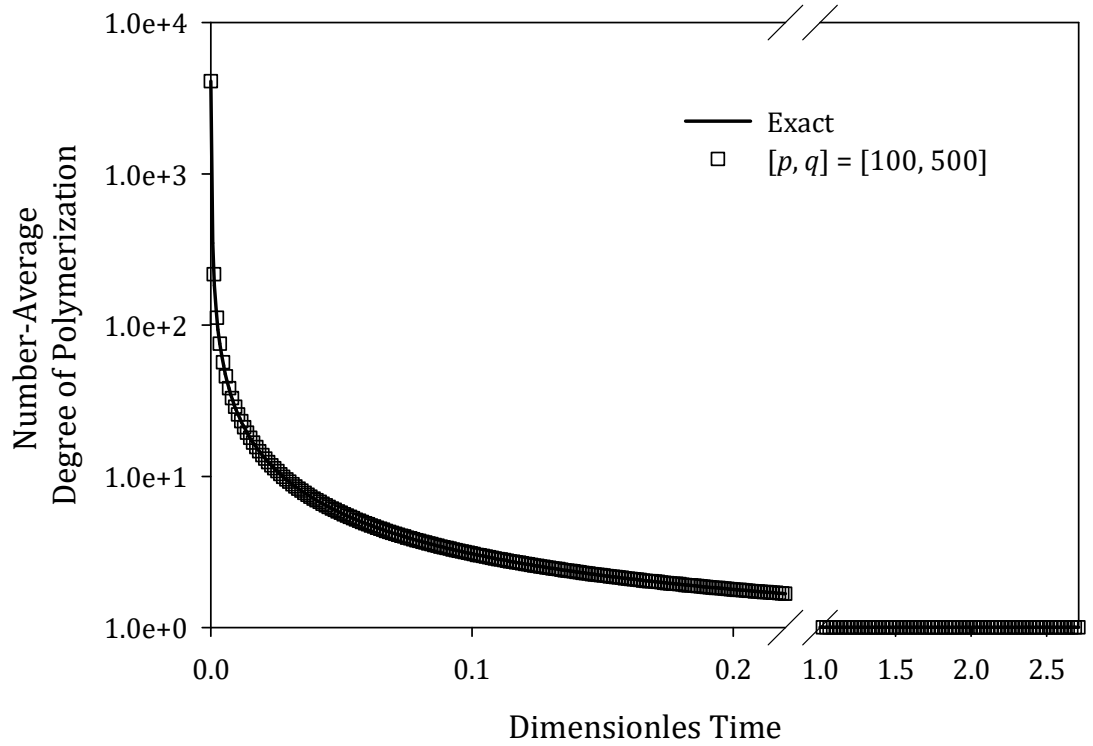


Figure 4.19: Transient of the number-average DP using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. Here, $r = 1.0109$ for $[p, q] = [100, 500]$. The dimensionless time is normalized against the time required for 99% monomer production.

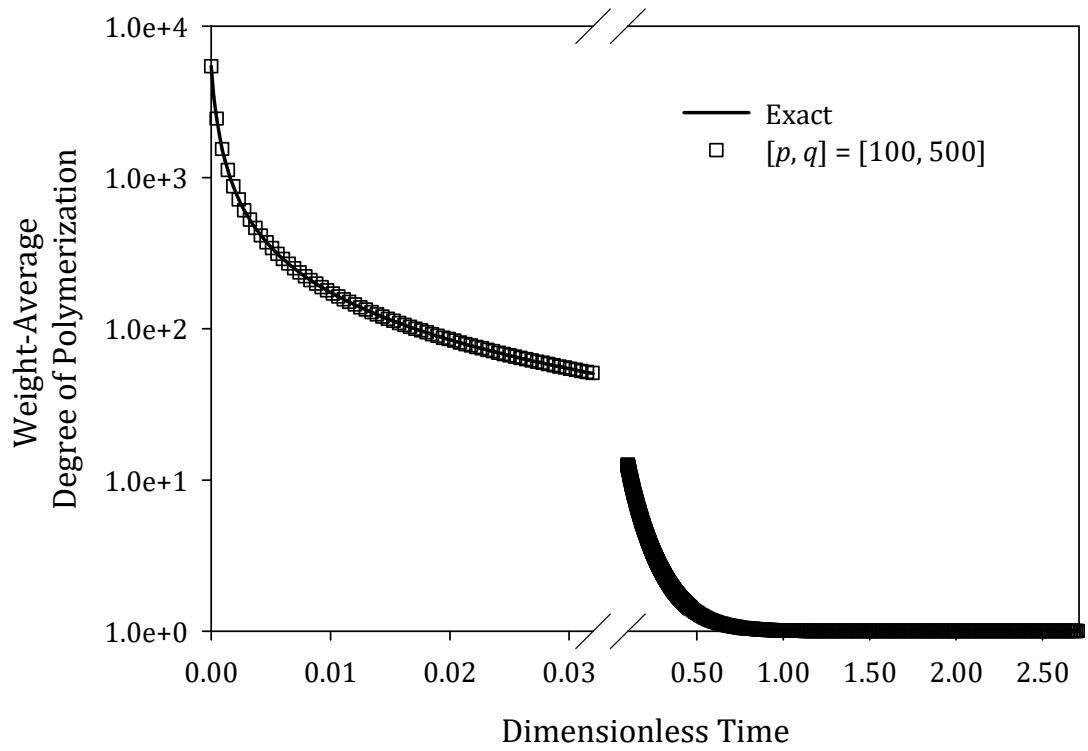


Figure 4.20: Transient of the weight-average DP using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. Here, $r = 1.0109$ for $[p, q] = [100, 500]$. The dimensionless time is normalized against the time required for 99% monomer production.

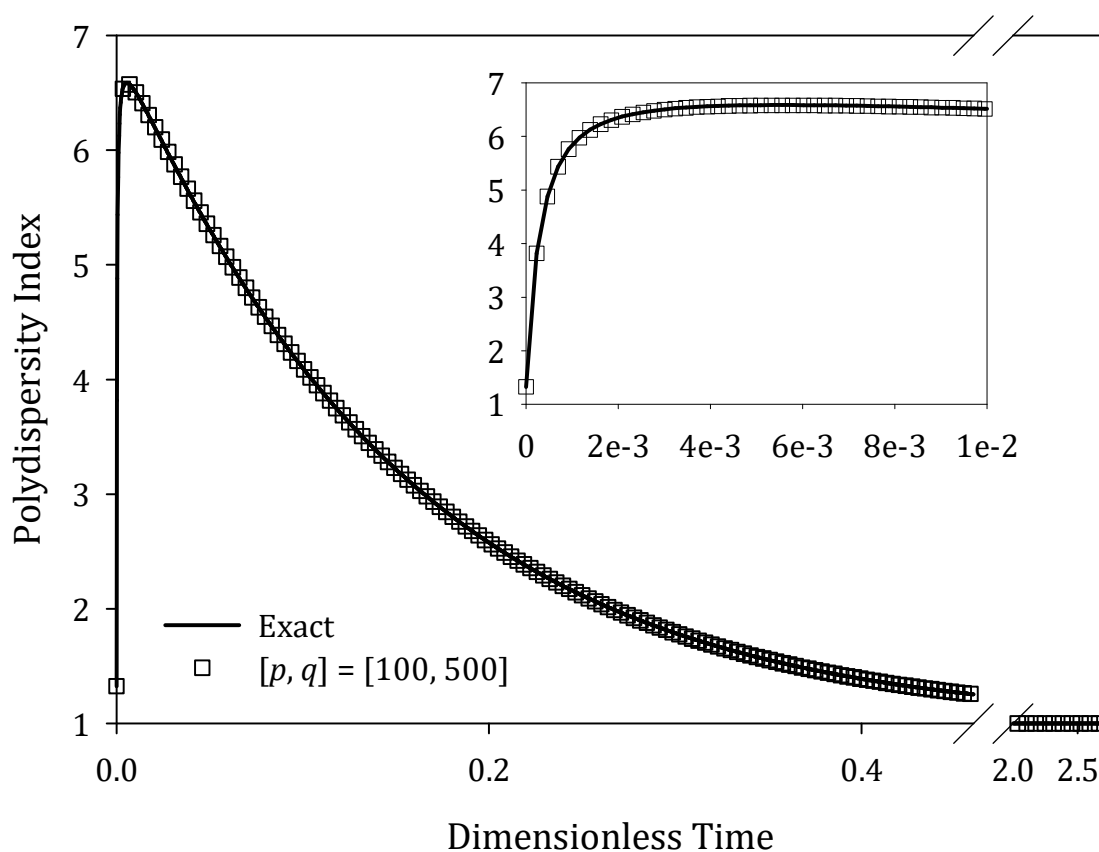


Figure 4.21: Transient of the polydispersity index using the exact and the fixed pivot (FP) solutions for simultaneous random and chain-end scissions. Here $[p, q] = [100, 500]$ and $r = 1.0109$. The dimensionless time is normalized against the time required for 99% monomer production.

In all the simulations performed in this section, even with the markedly reduced time incurred for computing the exact solutions, the FP solution on average took only less than 3.5% (≈ 18.5 s) of the time required for computing the exact solution (≈ 530 s). Further test was performed by increasing the scale of the problem to $N = 10 \times 22496$, with $p+q = 2000$ ($< 1\%$ of N) in this case. The molar concentration density at $\theta = 1.23$, i.e. the time further beyond 99% formation of monomer, is shown in Figure 4.22. Once again, the results show that the FP technique predicted the exact solution accurately even at such a late phase of the reaction. On the workstation C, the computational time required by the solution technique was merely 1.2% (≈ 9 mins) of the duration (≈ 13 hrs) required by the

4.3 Observations of the FP Method for Solving Random Scission

In Section 4.1.5, the observations of the FP method in solving chain-end scission had been presented. Although Section 4.2 mainly concerns the solution of simultaneously occurring random and chain-end scissions, in order to observe the effect of having random scission as part of the problem formulation, the following discussion will be dedicated to studying only the random scission. From the FP equation for pure random scission [Eq. (4.59)], the rate of change in the concentration of the polymer at the i -th pivot can be re-written as follows:

$$\frac{dC_i(t)}{dt} = \sum_{j=i+1}^{p+q} \left[\frac{x_{i+1} - x_i}{x_j - 1} + \frac{x_i - x_{i-1}}{x_j - 1} \right] k_j^\alpha C_j(t) - \left[\frac{x_{i-1} - 1}{x_i - 1} \right] k_i^\alpha C_i(t) \quad (4.68)$$

As shown in Section B.2 of Appendix B, this equation can be simplified to the following form:

$$\frac{dc_i(t)}{dt} = x_{p+1} \sum_{j=i+1}^{p+q} \left[\frac{r^{j-p} - r^{j-p-2}}{x_j - 1} \right] k_j^\alpha c_j(t) - \left[\frac{r^{i-1-(p+1)} x_{p+1} - 1}{r^{i-(p+1)} x_{p+1} - 1} \right] k_i^\alpha c_i(t) \quad (4.69)$$

From Eq. (4.69), in the limit where r approaches that of an arithmetic mesh (with the pivots being spaced 1 unit apart), the equation should approximate the exact form. However, as the formulae of $x_{i+1} = rx_i$ was used in the formulation, the limit of $r = 1$ (which was used in Section 4.1.5) which will result in the erroneous expression of ‘ $x_{i+1} = x_i$ ’ cannot be used here as this limit is applicable when only $r = \frac{x_i - x_{i-1}}{x_{i-1} - x_{i-2}}$ is used.

By using this ‘interval’ definition, r in such a case would be able to represent both geometric meshing ($r > 1$) as well as the exact meshing ($r = 1$). From Section B.2 of

Appendix B, Eq. (4.68) cannot be simplified to Eq. (4.69) by invoking $r = \frac{x_i - x_{i-1}}{x_{i-1} - x_{i-2}}$

alone, but that the definition of $x_{i+1} = rx_i$ is also necessary. Since it is not possible here

to avoid using the conventional definition of r , i.e. $x_{i+1} = rx_i$ where $r > 1$, a new limit which approximates the exact solution is required. Since the distance between $x_{p+2} - x_{p+1}$ was established to be at least $v_m = 1$ unit apart in Section 4.1.3, it follows that $x_{p+2} - x_{p+1} \geq v_m$, leading to the following inequality for r :

$$r \geq 1 + \frac{1}{p+1} \quad (4.70)$$

Thus, $r = 1 + 1/(p+1)$ could serve as an approximation to arithmetic meshing. This can be validated by the following proof. If the new limit is capable of approximating the arithmetic mesh, substituting this r into $x_{p+q} = r^{q-1}x_{p+1}$ should yield $p+q$. Let $r = 1 + u$ where $u = (p+1)^{-1}$, it follows that:

$$\begin{aligned} x_{p+q} &= r^{q-1}x_{p+1} \\ &= (1+u)^{q-1}u^{-1} \quad (\because x_{p+1} = p+1) \\ &= \left[1 + (q-1)u + O(u^m)\right]u^{-1} \\ &= p+q + O(u^{m-1}) \\ &\approx p+q \quad (\because p \gg 1 \text{ and ignoring the } m \geq 2 \text{ terms}) \end{aligned} \quad (4.71)$$

In the above, the binomial theorem was used to expand the term $(1+u)^{q-1}$. From the proof, the new limit for r approximates well to the arithmetic mesh at large values of p .

Next, Eq. (4.69) can be re-written in the following manner:

$$\frac{dc_i(t)}{dt} = \sum_{j=i+1}^{p+q} f(r)k_j^\alpha c_j(t) - g(r)k_i^\alpha c_i(t) \quad (4.72)$$

where $f(r)$ and $g(r)$ are defined as:

$$f(r) = (p+1) \left[\frac{r^{j-p} - r^{j-p-2}}{x_j - 1} \right] \quad (4.73)$$

$$g(r) = \frac{r^{i-1-(p+1)}(p+1) - 1}{r^{i-(p+1)}(p+1) - 1} \quad (4.74)$$

Expanding $f(r)$ and $g(r)$ around the limit of arithmetic mesh by re-writing $r = (1+u) + \varepsilon = r^*$ where $u = (p+1)^{-1}$ and using the binomial theorem, the equations below can be obtained :

$$\begin{aligned}
f(r^*) &= f(\varepsilon) \\
&= \left[\frac{u^{-1}}{x_j - 1} \right] \left\{ [1 + (u + \varepsilon)]^{j-p} - [1 + (u + \varepsilon)]^{j-p-2} \right\} \\
&= \left[\frac{u^{-1}}{x_j - 1} \right] \left\{ \left[1 + (j-p)(u + \varepsilon) + O([u + \varepsilon]^m) \right] \right. \\
&\quad \left. - \left[1 + (j-p-2)(u + \varepsilon) + O([u + \varepsilon]^m) \right] \right\} \\
&= \left[\frac{u^{-1}}{x_j - 1} \right] \left\{ 2(u + \varepsilon) + O([u + \varepsilon]^m) \right\}
\end{aligned} \tag{4.75}$$

$$\begin{aligned}
g(r^*) &= g(\varepsilon) \\
&= \frac{[1 + (u + \varepsilon)]^{i-p-2} u^{-1} - 1}{[1 + (u + \varepsilon)]^{i-p-1} u^{-1} - 1} \\
&= \frac{[1 + (i-p-2)(u + \varepsilon) + O([u + \varepsilon]^m)] u^{-1} - 1}{[1 + (i-p-1)(u + \varepsilon) + O([u + \varepsilon]^m)] u^{-1} - 1} \\
&= \frac{p + u^{-1}(i-p-2)(u + \varepsilon) + u^{-1}O([u + \varepsilon]^m)}{p + u^{-1}(i-p-1)(u + \varepsilon) + u^{-1}O([u + \varepsilon]^m)}
\end{aligned} \tag{4.76}$$

In the above, $m \geq 2$.

Finally, Eqs. (4.72) - (4.74) can be re-written in the following form by using the results of Eqs. (4.75) - (4.76):

$$\frac{dc_i(t)}{dt} = \sum_{j=i+1}^{p+q} f(\varepsilon) k_j^\alpha c_j(t) - g(\varepsilon) k_i^\alpha c_i(t) \tag{4.77}$$

$$f(\varepsilon) = \left[\frac{u^{-1}}{x_j - 1} \right] \left\{ 2(u + \varepsilon) + O([u + \varepsilon]^m) \right\}; \quad m \geq 2; \quad u = (p+1)^{-1} \tag{4.78}$$

$$g(\varepsilon) = \frac{p + u^{-1}(i-p-2)(u + \varepsilon) + u^{-1}O([u + \varepsilon]^m)}{p + u^{-1}(i-p-1)(u + \varepsilon) + u^{-1}O([u + \varepsilon]^m)}; \quad m \geq 2; \quad u = (p+1)^{-1} \tag{4.79}$$

Three observations could be gleaned from Eqs. (4.77) - (4.79) regarding the use of the FP technique in solving PBEs which involve random scission. First, when $\varepsilon \rightarrow 0$ (approaching the arithmetic mesh), and ignoring the higher order terms involving $m \geq 2$ because $p \gg 1$, the equations simplify to:

$$\frac{dc_i(t)}{dt} = \sum_{j=i+1}^N \left[\frac{2}{x_j - 1} \right] k_j^\alpha c_j(t) - \left[\frac{i-2}{i-1} \right] k_i^\alpha c_i(t) \quad (4.80)$$

At large values of i , $(i-2)/(i-1) \approx 1$ and thus Eq. (4.80) collapses to the exact solution for pure random scission. Therefore, the departure of r from that which approximates the arithmetic mesh could then be seen as the extent of deviation from the exact solution. Fortunately, the values of r resulting from the proposed meshing strategy in Section 4.1.3 is always close to the limit of the arithmetic mesh because $q-1$ in $r = \left[x_{p+q}/x_{p+1} \right]^{\frac{1}{q-1}}$ is usually large. For instance, $r = 1.0109$ for the results presented in Figures. 4.16 – 4.21 for $[p, q] = [100, 500]$ and the limit to arithmetic mesh is $r = 1 + 1/(p+1) = 1.0099$.

Second, at small values of i , the FP technique tends to under-estimate the death term, i.e. $(i-2)/(i-1) < 1$, even when r approximates that of the arithmetic mesh. This is the inherent nature of the FP approximation for random scission, where the birth term tends to receive more than what the death term is giving out, resulting in an over-prediction in the final mass of the population. This over-prediction is more so evident when r departs from the limit of arithmetic mesh. As proven in Section B.3 of Appendix B, this is because $g(\varepsilon)$ [Eq. (4.79)] decreases further from unity and $f(\varepsilon)$ [Eq. (4.78)] increases further from $2/(x_j - 1)$ as r deviates from the limit of arithmetic mesh.

Finally, as alluded to previously, the value of p has an effect on the accuracy of the solution. Referring to Eqs. (4.77) - (4.79), regardless of the extent of the departure of r

from the arithmetic mesh, a large value of p generally improves the accuracy of the solution by minimizing $(u + \varepsilon)$. To further illustrate, the FP technique for pure random scission was implemented on the same example problem by Stickel and Griggs (2012) described in Section 4.1.5 with an initial distribution as shown in Eq. (3.4) with c_{in} , ω and κ given as 100, 400 and 1.5 respectively. Example codes for this simulation using the FP and the exact solution are given in Sections B.4 – B.5 of Appendix B. Figure 4.23 shows the results of the simulation. Continuing the previous practice, the pivots were spaced evenly in the continuous region with $p+q = 555$. The common difference for $[p, q] = [10, 545]$ is 1.5 whereas that for $[p, q] = [100, 455]$ is 1.6. Such closely spaced pivots are employed here to highlight the effect of p on the accuracy of the solution. From the results, within a constant $p+q$, a larger value of p greatly improves the solution accuracy. As expected, at a low value of $p = 10$, the over-predictions observed were significant. For DP2, increasing the value of p reduced ε_g by two orders of magnitude, i.e. from $\sim O(10^{-1})$ to $\sim O(10^{-3})$, whereas for DP7, the reduction achieved was one order of magnitude, i.e. from $\sim O(10^{-2})$ to $\sim O(10^{-3})$. Although DP2 – DP7 do not belong to the continuous region, they nonetheless still contain the birth terms of which their departure from the exact solution is as elaborated in the foregoing discussion. The results here are important in that by merely increasing the number of pivots in the discrete region while conserving the total number of pivots (and hence the computational burden incurred), the accuracy of the solution can be significantly improved.

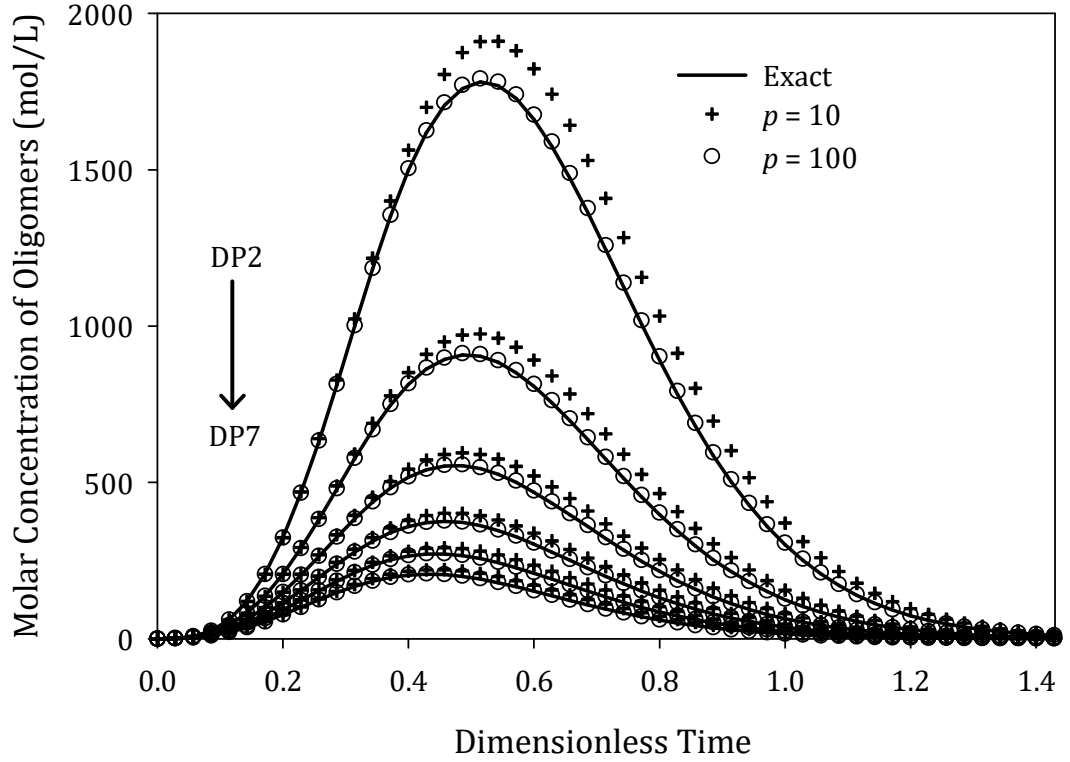


Figure 4.23: Temporal evolution of molar concentration of the DP2 – DP7 oligomers using the fixed pivot (FP) and the exact solutions for random scission. The dimensionless time $\theta = 1$ corresponds to the time when more than 99% monomer had been formed. Here, $[p, q] = [10, 545]$ and $[p, q] = [100, 455]$ were used where p and q are the number of pivots for the discrete and continuous region respectively.

Based on the above, when random scission is involved, one additional point to consider during meshing is that the value of p could be used to improve the accuracy of the solution. From Eqs. (4.77) - (4.79), a rule-of-thumb guide for choosing the value for p can be established by inspecting $f(\varepsilon)$ at $\varepsilon = 0$. The rationale for doing so is that should the deviation from the exact solution at this best case be large, that particular value of p is inappropriate. In addition, $g(\varepsilon)$ is not examined here because at the discrete region, the death term is exactly $-k_i^\alpha c_i(t)$ and thus only $f(\varepsilon)$ projects the effect of meshing at the continuous region on the smaller oligomers of interest in the discrete region. At the

limit of the arithmetic mesh, $\lim_{\varepsilon \rightarrow 0} f(\varepsilon) = \frac{2}{x_j - 1} + \frac{1}{x_j - 1} O(u^{m-1})$ where $m \geq 2$. Thus, the source of the over-prediction originates from the terms involving $m \geq 2$ because the first order term is the exact solution. Using only the $m = 2$ term and given $u = (p+1)^{-1}$, this source of over-prediction in the formulation (ε_{pred}) is of the order of u :

$$\varepsilon_{pred} \sim \frac{1}{p+1} \quad (4.81)$$

Using Eq. (4.81), the value of p can be chosen such that the value of ε_{pred} is small. For example, throughout Figures 4.16 – 4.21, the value of $\varepsilon_{pred} \sim O(10^{-3})$ for $p = 100$. Practically, within the constraint of $q < 1 + \left\{ \ln \left[N / (p + v_m) \right] / \ln \left[1 + v_m / (p + v_m) \right] \right\}$ established in Section 4.1.3, typically values of $\varepsilon_{pred} \leq O(10^{-2})$, i.e. $p = 50 \sim 100$, is a good starting point. Assuming that values of N and v_m are typically known, the revised meshing guidelines are:

- (a) Choose $p+q \sim O(10^{-2}N)$. Then choose the value of p such that the constraints of

$$q < 1 + \left\{ \ln \left[N / (p + v_m) \right] / \ln \left[1 + v_m / (p + v_m) \right] \right\} \quad \text{and} \quad (p+1)^{-1} \leq O(10^{-2}) \quad \text{are fulfilled.}$$

There is a lower limit for p below which q becomes too large, causing pivots to be spaced closer than v_m . To facilitate this, Figure 4.13 in Section 4.1.6 can be conveniently used to ensure that $p \geq p_{min}$.

- (b) For values of $p+q$ and N not covered in Figure 4.13, first choose p such that

$$(p+1)^{-1} \leq O(10^{-2}) \quad \text{and calculate } q \text{ using the constraint in (a). If the resulting } p+q \sim$$

$O(10^{-2}N)$, proceed to solve the system of differential equations. Otherwise, select a different value of p and repeat the procedure until the desired $p+q$ is found.

4.3 Concluding Remarks

Chain-end and random scissions as modeled by PBM can be efficiently and accurately solved with the FP technique via a judicious choice of mesh parameters to partition the computational domain into a discrete and a continuous region. For chain-end scission, except for unnaturally narrow distributions of the polymer sizes, good approximation was obtained with the number of equations not exceeding a few percent of the exact case, at a fraction of the time required for the exact case. When deployed together with random scission, except for the inherent tendency to over-predict at small number of pivots in the discrete region, similar performance was demonstrated by the FP technique when benchmarked against the exact solution which was solved at an improved efficiency. The formulation here readily provides the concentration profiles of the oligomers, a convenience that is particularly significant when only these are analyzable by existing experimental techniques. As far as enzymatic depolymerization is concerned, the current formulation covers the commonly encountered depolymerization phenomena, a tool that is indispensable for the study of related systems.

CHAPTER 5 : INTERLINKED POPULATION BALANCE AND CYBERNETIC MODELLING FRAMEWORK

In the previous chapter, the numerical techniques necessary for solving PBEs had been fully established and shown to be very appropriate for simulating the profiles of various oligomers commonly found in polymeric fermentation systems. This is an important groundwork for the materials which will be presented in this chapter, which is to lay out the general framework for interlinking the PBM and the CM methodologies. Section 5.1 is devoted to such a purpose, followed by a demonstration of the capability of the resulting framework by two case studies in Sections 5.2 and 5.3 respectively.

5.1 Theoretical Framework

The general idea of the linkage between the PBM and the CM components had been elaborated in Figure 1.1 of Section 1.1. In this section, the theoretical framework for the interlinked model will be established. For compactness, equations presented in this section are written as vector equations. Unless specifically denoted as constants, all variables are functions of time. For Sections 5.1.1 – 5.1.3, the equations presented are general and do not refer to a particular mode of enzymatic scission. For the case studies, however, the convention of using the superscripts ' α ' and ' γ ' to distinguish between the quantities associated with enzymes which exhibit random and chain-end scissions respectively is adopted.

5.1.1 Population Balance Modelling for Enzymatic Scission

In general, an enzymatic depolymerization process may contain multiple enzymes which can act on multiple substrates in the broth. Typical of all enzymatic processes, the formation of enzyme-substrate complexes precedes the enzymatic scission. As polymers are generally long chain macromolecules, a possibility exists where multiple enzymes bind with different parts of the chain (Griggs et al., 2012a), resulting in numerous different complexes in the reaction broth. To represent the molar concentrations of different species in the broth, all substrate-related terms (including the complexes) are collected in a vector \mathcal{S} . The molar concentrations of all the free (or unbounded) enzymes are collated in a vector E . This is further clarified in Table 5.1 for one possible scenario of enzyme complex formation between two substrates and two enzymes.

Table 5.1: The components of \mathcal{S} and E for one possible scenario of enzyme complex formation between two substrates of different chain length and two different enzymes. The number of enzymes and substrates are generally not restricted to two.

\mathcal{S}	E

For a broth containing a polymer with a maximum DP equals N and extracellular depolymerases, the PBEs describing the temporal evolution of all the polymer species, written together with the enzyme balance equation can be expressed in the following form:

$$\frac{d\mathcal{S}}{dt} = \underbrace{f(E, \mathcal{S}, k)}_{\text{Enzymatic Generation}} - \underbrace{g(E, \mathcal{S}, k)}_{\text{Enzymatic Scission}} \quad (5.1)$$

$$E = E(0) - G; \quad G = T\mathcal{S} \quad (5.2)$$

Here, k is the vector of rate constants, $E(0)$ is the vector of the initial enzyme loading, and G is the sum of the molar concentrations of all of the enzyme complexes. The matrix T , which is stationary in time, extracts enzyme-bounded substrates from \mathcal{S} .

The generation term in Eq. (5.1) represents the enzymatic scission of larger polymers with $DP > i$ to form $DP = i$, while the scission term caters to the scission of the current polymer with $DP = i$ to smaller ones with $DP < i$. As the value of N (i.e. DP of the largest polymer) can be very large (which is not uncommon for natural polymers), solving Eqs. (5.1) - (5.2) can be computationally expensive. Amongst the many methods to solve PBEs (D. Ramkrishna, 2000), the Fixed Pivot (FP) technique established in Chapter 4 was used in the current framework to drastically reduce the number of ODEs to be solved. Recalling briefly, the strategy involves the partitioning of the DP domain into a discrete region for the smaller oligomers and a continuous region for the larger oligomers. This reduction is done by replacing the first term on the RHS of Eq. (5.1) with $f(E, \mathcal{S}, k, n)$ where constants n_{ij} represents the fractional allocation of polymers splitting from $DP = j$ into i by conserving the zeroth and the first moments of the distribution. The general expression for n_{ij} is given as:

$$n_{ij} = \int_{x_i}^{x_{i+1}} \left[\frac{x_{i+1} - v}{x_{i+1} - x_i} \right] b(v, x_j) dv + \int_{x_{i-1}}^{x_i} \left[\frac{v - x_{i-1}}{x_i - x_{i-1}} \right] b(v, x_j) dv \quad (5.3)$$

where v is the continuous DP, x is the pivot DP and $b(v, x_j)$ is the stoichiometric kernel relating the formation of polymers with DPs of v and $x_j - v$ from x_j (D. Ramkrishna, 2000). On applying this approximation, the sizes of \mathcal{S} and \mathbf{T} are reduced. Useful properties of the polymer distribution which can be extracted from the PBEs are the moments of the polymer distribution, with the ξ -th moment given as:

$$\mathcal{S}^{(\xi)} = \|\mathbf{H}\mathbf{L}\|_1; \quad \mathbf{H} = \begin{bmatrix} x_1^\xi & x_2^\xi & \cdots & x_{p+q}^\xi \end{bmatrix}^T; \quad \mathbf{L} = \mathbf{F}\mathcal{S} \quad (5.4)$$

The matrix \mathbf{F} (stationary in time) operates on \mathcal{S} such that the i -th element of vector \mathbf{L} contains the sum of all species with $DP = i$. Here, p and q are the meshing parameters representing the number of pivots in the discrete and continuous regions respectively (cf. Section 4.1.3). Since mass should be conserved, the first moment of the distribution ($\xi =$

1) is a constant. The PBE presented here can be used to model a pure enzymatic system where the involvement of microbes is not present. Example implementations for specific mechanisms of scission are given in the Appendix D. In the presence of microbes, an additional set of equations for microbial kinetics is required, as discussed next.

5.1.2 Cybernetic Modelling

As the CM framework developed by Ramkrishna and co-workers (D. Ramkrishna & Song, 2012; Song et al., 2009) had been presented in Section 2.3, it shall not be re-elaborated here. Nonetheless, the equations are re-produced here to facilitate the discussions in the subsequent sections. Briefly the CM equations are:

$$\frac{1}{X} \frac{d\mathcal{N}}{dt} = \mathbf{W}_{\mathcal{N}} \mathbf{r}(V, \mathbf{e}, \mathcal{N}) \quad (5.5)$$

$$\frac{1}{X} \frac{d\mathcal{P}}{dt} = \mathbf{W}_{\mathcal{P}} \mathbf{r}(V, \mathbf{e}, \mathcal{N}) \quad (5.6)$$

$$\frac{d\phi}{dt} = \mathbf{W}_{\phi} \mathbf{r}(V, \mathbf{e}, \mathcal{N}) - \mu \phi \quad (5.7)$$

$$\frac{1}{X} \frac{dX}{dt} = \mu \quad (5.8)$$

$$\frac{d\mathbf{e}}{dt} = \mathbf{p}_e + \mathbf{D}(U) \mathbf{r}_e - [\mathbf{D}(\beta) + \mu \mathbf{I}] \mathbf{e} \quad (5.9)$$

where X is the mass concentration of biomass, μ is the total specific growth rate, and \mathcal{N} is the concentrations of extracellular substrates, \mathcal{P} is the concentration of extracellular products, and ϕ is the concentration of intracellular metabolites. The matrices $\mathbf{W}_{\mathcal{N}}$, $\mathbf{W}_{\mathcal{P}}$ and \mathbf{W}_{ϕ} contain the stoichiometry of the reactions. The symbol \mathbf{r} is the regulated fluxes (or rates defined per unit of biomass), \mathbf{p}_e is the constitutive enzyme synthesis rates, \mathbf{r}_e is the inducible enzyme synthesis rates, β is the enzyme degradation rate constants and \mathbf{e}

is the enzyme levels. Unless dimensionless, units of concentration related terms may be expressed either on a molar or a mass basis wherever appropriate. The matrix \mathbf{I} is the identity matrix and the operator $\mathbf{D}(\bullet)$ converts the input vector into a diagonal matrix such that:

$$\begin{aligned} \mathbf{D}(\boldsymbol{\beta}) &= \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \beta_{n_e} \end{bmatrix}; \quad \mathbf{D}(\mathbf{U}) = \begin{bmatrix} U_1 & 0 & \cdots & 0 \\ 0 & U_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_{n_e} \end{bmatrix}; \\ \mathbf{D}(\mathbf{V}) &= \begin{bmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & V_{n_e} \end{bmatrix} \end{aligned} \quad (5.10)$$

where n_e is the size of \boldsymbol{e} . The general form of \mathbf{U} and \mathbf{V} can be written as (Song et al., 2009):

$$\mathbf{U} = \frac{\mathbf{R}^+}{\sum_k R_k^+}; \quad \|\mathbf{U}\|_1 = 1, U_k \geq 0 \quad (5.11)$$

$$\mathbf{V} = \frac{\mathbf{R}^+}{\max(\mathbf{R}^+)}; \quad V_k \leq 1 \quad (5.12)$$

where $\mathbf{R}^+ = [R_1^+ \quad R_2^+ \quad \cdots \quad R_k^+]^T$ with $R_k^+ = \max(R_k, 0)$ and R_k is the return on investment from the k -th alternative. To further clarify the construction of the CM equations, an example is given in Section E.1 of Appendix E. As alluded to previously in Section 2.3.1, although the CM equations given above can be used to predict the effect of multiple substrates on microbial growth, the existing framework does not cater explicitly to systems where the excretion of the extracellular depolymerases is required to break large polymers into consumable substrates.

5.1.3 Interlinked Population Balance and Cybernetic Framework

As alluded to previously in Section 1.1, to connect the population balance and the cybernetic framework, the PBE must be modified to include a term for substrate uptake by the microbes. This is done by rewriting the PBE in Section 5.1.1:

$$\frac{d\mathcal{S}}{dt} = \underbrace{f(E, \mathcal{S}, k, n)}_{\text{Enzymatic Generation}} - \underbrace{g(E, \mathcal{S}, k)}_{\text{Enzymatic Scission}} - \underbrace{W_{\mathcal{N}} r(V, e, \mathcal{N}) X}_{\text{Microbial Uptake}} \quad (5.13)$$

For simplicity, it is assumed that $\mathcal{N} \subset \mathcal{S}$, i.e. all the carbon and energy sources are part of the polymeric substrates. The free extracellular depolymerization enzymes (E) is now dependent on the output of CM. This is the second critical link which in the past was barely addressed. In several studies which employed the simplistic M-M kinetics for hydrolysis and the CM framework for microbial growth (Altintas et al., 2002; Ochoa et al., 2007), the rate of enzymatic scission was assumed to be a function of the dimensionless intrinsic enzyme level. This obscures the underlying molecular nature of the enzymatic action, e.g. the binding of one extracellular enzyme molecule with one substrate molecule. Accounting for the enzyme-substrate interaction on a molar basis is more convenient for the establishment of the PBM and CM linkage because the PBM framework concerns the temporal evolution of the number of polymer molecules.

Extracellular enzymes, like their intracellular counterparts, are subjected to cellular regulation according to the mechanism of induction/repression (Kobayashi & Nakamura, 2003, 2004; Nakamura et al., 1997), of which their synthesis first occur in the biotic phase prior to being excreted into the abiotic phase. As the synthesis of both intra- and extra-cellular enzymes are the result of an optimal allocation of a common pool of critical resources, construction of their respective enzyme synthesis equation should adhere to identical CM principles. For simplicity, the enzymes produced within the cells are

assumed to be excreted instantaneously into the external broth and do not accumulate appreciably within the cells, thereby allowing the contention with only their external concentrations. The rate equations for both intra- and extracellular enzyme synthesis are given respectively as:

$$\frac{de}{dt} = \rho_e + D(U_e)r_e - [D(\beta_e) + \mu I]e \quad (5.14)$$

$$\frac{dE^0}{dt} = \rho_E X + D(U_E)r_E X - D(\beta_E)E^0 \quad (5.15)$$

where the subscript ‘E’ is used to distinguish between the parameters for the extracellular enzymes from that of the intracellular enzymes (subscript ‘e’). The form of Eq. (5.15) is derived in Section F.1 of Appendix F. Here, E^0 represents the total molar concentrations of extracellular enzymes at a particular instance. As only the free enzymes are capable of forming new complexes with the substrates, the molar concentration of the free extracellular enzymes must be obtained from enzyme balance:

$$E = E^0 - G; \quad G = T\mathcal{S} \quad (5.16)$$

Note that the total enzyme concentration in Eq. (5.16) evolves with time while for pure enzymatic scission, the total enzyme concentration (i.e. the initial enzyme loading) in Eq. (5.2) is a constant (normally temporal degradation is ignored).

Differing from the traditional CM which makes no distinction between the intracellular and extracellular enzymes, in Eqs. (5.14) and (5.15) the cybernetic variable U must account also for the induction/repression of extracellular enzymes. U is partitioned into intracellular (U_e) and extracellular (U_E) parts, i.e. $U = [U_e \quad U_E]^T$, given as:

$$U_e = \frac{R_e^+}{\sum_k R_k^+}; \quad U_E = \frac{R_E^+}{\sum_k R_k^+}; \quad \|U\|_1 = 1 \quad (5.17)$$

where the return on investment $\mathbf{R}^+ = \begin{bmatrix} \mathbf{R}_e^+ & \mathbf{R}_E^+ \end{bmatrix}^T$. This formulates the excretion of the extracellular enzymes as part of the cellular response for ensuring its survival, which in tandem to the synthesis of intracellular metabolic enzymes, requires the maximization of a common metabolic objective function.

On the other hand, it is reasonable to assume that the activities of the depolymerization enzymes, once excreted into the external broth, are not regulated by the microbes. Thus only the intracellular metabolic enzymes are regulated via the cybernetic variable V :

$$V = \frac{\mathbf{R}_e^+}{\max(\mathbf{R}_e^+)} \quad (5.18)$$

The remaining equations for product formation, accumulation of intracellular metabolites, and biomass growth remain unchanged as compared to the standard CM formulation:

$$\frac{1}{X} \frac{d\mathcal{P}}{dt} = \mathbf{W}_{\mathcal{P}} \mathbf{r}(V, \mathbf{e}, \mathcal{N}) \quad (5.19)$$

$$\frac{d\phi}{dt} = \mathbf{W}_{\phi} \mathbf{r}(V, \mathbf{e}, \mathcal{N}) - \mu \phi \quad (5.20)$$

$$\frac{1}{X} \frac{dX}{dt} = \mu \quad (5.21)$$

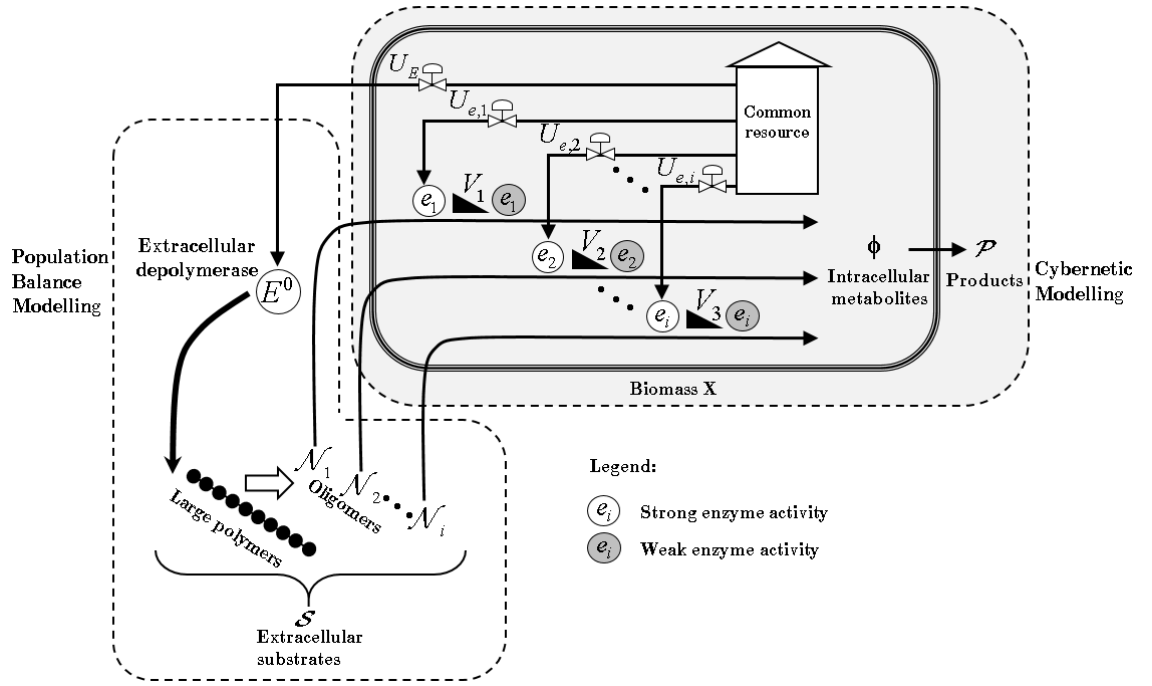


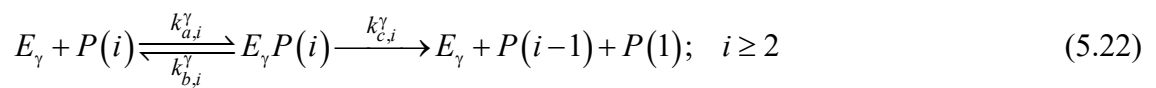
Figure 5.1: Simplified diagram illustrating the different components of the interlinked PBM and CM framework where the shaded region (excluding the excretion of extracellular depolymerase) is the standard CM framework. In this illustration, for simplicity the cell is assumed to excrete only one form of depolymerase for the interlinked framework.

Equations (5.13) - (5.21) represent the general framework for interlinking the PBM and the CM equations along with the pictorial illustration shown in Fig. 5.1. Regardless of their respective variants, as long as the withdrawal of nutrients by the microbes and the excretion of extracellular enzymes are pertinent, the framework is applicable. The applicability of the framework will be demonstrated through two case studies in the following sections. The first case study concerns the growth of a glucoamylase producing recombinant *S. cerevisiae* on starch while the second case study will showcase how the framework could be used to model the growth of another genetically engineered *S. cerevisiae* which could excrete two different forms of depolymerases simultaneously.

5.2 Case Study I: Growth of A Glucoamylase Producing Recombinant *S. cerevisiae* on Starch

5.2.1. Model Formulation

In the work of Nakamura et al. (1997), a recombinant yeast, i.e. *S. cerevisiae* SR93, was constructed by integrating a glucoamylase producing gene (*STAI*) originating from *Saccharomyces diastaticus* into the chromosome of *S. cerevisiae* SH1089. In the presence of starch, the yeast produces glucoamylase which releases successive glucose units from the non-reducing ends of starch molecules. Although starch consists of a mixture of linear (amylose) and branched (amylopectin) components, Chang et al. (2002) pointed out that the effect of molecular branching on the overall rate of hydrolysis is not significant. This was validated by experimental observations that starch in general contains a low proportion of branching, i.e. approximately 5% within a starch molecule, and that there is minor difference in the rates of hydrolysis between the branched and unbranched molecules (Dean III & Rollings, 1992). As the variations in the kinetics of hydrolysis caused by the branching effect can also be accommodated by the adjustment of the kinetic parameters without changing the reaction mechanisms, for simplicity starch was modelled here as a linear population of polymers. The hydrolysis mechanism is given as:



where $P(i)$ is the polymer with $DP = i$. As glucoamylase does not exhibit single chain attack or multiple attack patterns (Robyt, 2009), upon successful bond scission, $P(i-1)$ and $P(1)$ leave the enzyme active site, resulting in the recovery of the free enzyme. From the mechanism, the reduction of the number of vacant active sites of glucoamylase by

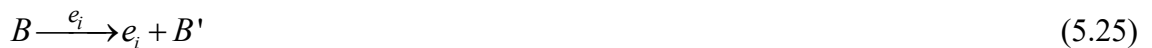
$P(1)$ (which traditionally is interpreted as the inhibition of enzyme action by glucose) is represented by Eq. (5.23). For the inhibition of enzyme action by starch, it was reported in the study by Miranda and Murado (1991) that this is the result of mass transfer limitation in the viscous broth. For simplicity, this was not considered here.

Furthermore, it is known that yeast does not consume starch directly, but that only the small sugars resulting from the breakdown of starch are consumed to produce ethanol under anaerobic conditions. Among the starch hydrolysates, yeast utilizes glucose (DP1), maltose (DP2), and maltotriose (DP3) in this approximate sequence (Batistote, Cruz, & Ernandes, 2006). Although yeast could consume ethanol during low sugar condition (Thomson et al., 2005), this was not observed in the study by Nakamura et al. (1997). For the purpose of illustrating the framework, the CM framework of Kompala and Ramkrishna (1986) is adopted here. Their model, although simplistic from the viewpoint of microbial metabolism, demonstrated the cybernetic principles and became the foundation for the subsequent development of more sophisticated variants. Metabolic burden, defined as the hurting of the normal metabolic functioning of the host cell by genetic manipulation (Glick, 1995), was not considered in the following formulation for simplicity. According to their framework, the assimilation of substrate $P(i)$ ($i = 1$ to 3) is catalysed by a key intracellular metabolic enzyme e_i :



where B is the biomass and Y_i is the yield of biomass from the utilization of substrate i .

The key enzyme e_i is induced in the presence of $P(i)$ according to:



In the above, B' is the biomass excluding the key enzyme. The accumulation of biomass (X , [g-DW/L]), which is a result of the utilization of the three sugars, can be written as:

$$\frac{dX}{dt} = \mu X; \quad \mu = \sum_{i=1}^3 r_{X,i} V_i \quad (5.26)$$

The specific rate of biomass growth on sugar i ($r_{X,i}$, [g/g-DW/h]) is dependent on the total amount of sugar i in the broth and is known to be inhibited by the presence of ethanol (Kobayashi & Nakamura, 2003, 2004; Nakamura et al., 1997):

$$r_{X,i} = \frac{\mu_{\max,i} \left[e_i / e_{\max,i} \right] \left\{ M_i \left[C_i + C_{B,i}^\gamma \right] \right\}}{\left\{ K_i + M_i \left[C_i + C_{B,i}^\gamma \right] \right\} \left\{ \left[1 + A / K_{EtOH} \right] \right\}}; \quad i = 1 \text{ to } 3 \quad (5.27)$$

In Eqs. (5.26) - (5.27), M_i [g/mol] ($i = 1, 2, 3$) is the molecular weight of glucose, maltose, and maltotriose respectively, C_i and $C_{B,i}^\gamma$ are the respective molar concentrations of unbounded and bounded sugar i , $\mu_{\max,i}$ [h^{-1}] and K_i [g/L] are the maximum specific growth rate and saturation constant of sugar i , A [g/L] is the mass concentration of ethanol, K_{EtOH} [g/L] is the constant of ethanol inhibition on growth, and $e_i / e_{\max,i}$ is the relative level of enzyme i . Note that the effect of inhibition by ethanol was assumed to be identical between the different sugars to avoid excessive parameterization.

As the production of ethanol is growth associated (Bailey & Ollis, 1986; Kobayashi & Nakamura, 2003, 2004; Shuler & Kargi, 2002), its accumulation is expressed as:

$$\frac{dA}{dt} = \sum_{i=1}^3 \left[\frac{Y_i^{EtOH}}{Y_i} \right] r_{X,i} V_i X \quad (5.28)$$

Here, Y_i [g/g] and Y_i^{EtOH} [g/g] are the biomass and ethanol yields from the utilization of sugar i . As the cybernetic model used here excluded the detailed pathways (Kompala & Ramkrishna, 1986), the intracellular metabolites (ϕ) were not modelled, thus omitting Eq. (5.20). The levels of intracellular metabolic enzymes for the assimilation of glucose, maltose, and maltotriose are given as:

$$\frac{d}{dt} \left[\frac{e_i}{e_{\max,i}} \right] = r_{e,i} U_{e,i} - \beta_{e,i} \left[\frac{e_i}{e_{\max,i}} \right] - \mu \left[\frac{e_i}{e_{\max,i}} \right]; \quad r_{e,i} = \frac{k_{e,i} \left\{ M_i [C_i + C_{B,i}^\gamma] \right\}}{K_{e,i} + M_i [C_i + C_{B,i}^\gamma]}; \quad i = 1 \text{ to } 3 \quad (5.29)$$

In this case, $e_i/e_{\max,i}$ is dimensionless where $e_{\max,i}$ [g/g-DW] is the maximum level of e_i [g/g-DW] and $k_{e,i}$ [h⁻¹] as well as $K_{e,i}$ [g/L] are the kinetic constants for enzyme synthesis. Here, the expression of $k_{e,i} = \mu_{\max,i} + \beta_{e,i}$ employed by several authors (Altintas et al., 2002; Kroumov et al., 2006; Ochoa et al., 2007) was not used following the reasons presented in Section E.2 of Appendix E.

For the critical link between the PBM and the CM equations, i.e. the temporal evolution of glucoamylase, it is known that glucoamylase production is induced by saccharides other than glucose, particularly starch (Kobayashi & Nakamura, 2003, 2004; Nakamura et al., 1997). Therefore, the rate of glucoamylase synthesis ($r_{E,\gamma}$, [g/g-DW/h]) is formulated as a function of the mass concentration of DP ≥ 2 . While DP = 2 does not represent starch, the model must be capable of handling the situation when only dimeric sugars are present, leading to synthesis of glucoamylase. Hence, the following expression is proposed followed by its corresponding enzyme balance equation:

$$\frac{dE_\gamma^0}{dt} = \frac{1}{M_\gamma} r_{E,\gamma} U_{E,\gamma} X - \beta_{E,\gamma} E_\gamma^0 - \beta_{EtOH} A^n E_\gamma^0; \quad r_{E,\gamma} = \frac{k_{E,\gamma} \sum_{j \geq 2}^{p+q} M_j [C_j + C_{B,j}^\gamma]}{K_{E,\gamma} + \sum_{j \geq 2}^{p+q} M_j [C_j + C_{B,j}^\gamma]} \quad (5.30)$$

$$E_\gamma = E_\gamma^0 - \sum_{j=1}^{p+q} C_{B,j}^\gamma \quad (5.31)$$

Here, M_γ [g/mol] refers to the molecular weight of glucoamylase, whereas $k_{E,\gamma}$ [h⁻¹] and $K_{E,\gamma}$ [g/L] are the kinetic constants of extracellular glucoamylase synthesis, and $\beta_{E,\gamma}$ [h⁻¹] is the deactivation rate constant. As the existing glucoamylase within the broth is inhibited by ethanol (Nakamura et al., 1997), an ethanol inhibition term was also added

where β_{EtOH} is the ethanol deactivation rate constant and η is the order of ethanol inhibition on glucoamylase production. For both the synthesis of intra- and extracellular enzymes, the constitutive term is omitted ($\rho_e = \mathbf{0}$ and $\rho_E = 0$) following Kompala and Ramkrishna (1986).

The cybernetic variables for controlling the synthesis of intra- and extracellular enzymes, used in Eqs. (5.29) - (5.30), are given as:

$$U_{e,i} = \frac{r_{X,i}X}{r_\gamma X + \sum_{j=1}^3 r_{X,j}X}; \quad U_{E,\gamma} = \frac{r_\gamma X}{r_\gamma X + \sum_{j=1}^3 r_{X,j}X}; \quad i = 1 \text{ to } 3 \quad (5.32)$$

Here, the return on investment for allocating cellular resources to synthesize e_i was taken to be the rate of biomass growth (Kompala & Ramkrishna, 1986) resulting from the utilization of sugar i . In addition, cellular resources are also allocated to synthesize glucoamylase. Although glucoamylase is not directly responsible for metabolizing DP ≥ 2 (primarily starch) in the biotic phase, in the absence of small sugars, glucoamylase is synthesized to produce consumable nutrients such as glucose from the larger saccharides. Therefore, yeast can be assumed to “grow” indirectly on starch by excreting glucoamylase, having a growth rate $r_\gamma X$ [g L⁻¹/h] similar to $r_{X,i}X$ [g L⁻¹/h] and is dependent on the concentrations of all species capable of inducing the production of glucoamylase:

$$r_\gamma = \frac{\mu_\gamma M_\gamma [E_\gamma^0/X] \sum_{j \geq 2}^{p+q} M_j [C_j + C_{B,j}^\gamma]}{\left\{ K_\gamma + \sum_{j \geq 2}^{p+q} M_j [C_j + C_{B,j}^\gamma] \right\} \left[1 + A/K_{EtOH} \right] \left\{ 1 + \left(M_1 [C_1 + C_{B,1}^\gamma] / K_I^\gamma \right) \right\}} \quad (5.33)$$

The form of the rate r_γ has a glucose inhibition term (K_I^γ [g/L] being the inhibition constant) and its form is further deliberated in Section F.2 of Appendix F. For the cybernetic variable which controls the activities of the enzymes (used in Eqs. (5.26)

and(5.28)), only the intracellular enzymes are regulated, hence r_γ is excluded from the expression:

$$V_i = \frac{r_{X,i}X}{\max(r_{X,1}X, r_{X,2}X, r_{X,3}X)}; \quad i = 1 \text{ to } 3 \quad (5.34)$$

Establishing the above, and referring to the hydrolysis mechanisms given in Eqs. (5.22) - (5.23), the material balance equations for glucose, maltose, maltotriose and those represented by $i = 4$ to $p+q$ are given as:

$$\frac{d}{dt} \begin{bmatrix} C_1 \\ C_{B,1}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,1}^\gamma C_{B,1}^\gamma + 2k_{c,2}^\gamma C_{B,2}^\gamma + \sum_{j=3}^{p+q} k_{c,j}^\gamma C_{B,j}^\gamma \\ k_{a,1}^\gamma E_\gamma C_1 \end{bmatrix} - \begin{bmatrix} k_{a,1}^\gamma E_\gamma C_1 \\ k_{b,1}^\gamma C_{B,1}^\gamma \end{bmatrix} - \frac{1}{M_1} \begin{bmatrix} \frac{1}{Y_1} \\ 0 \end{bmatrix} r_{X,1} V_1 X \quad (5.35)$$

$$\frac{d}{dt} \begin{bmatrix} C_2 \\ C_{B,2}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,2}^\gamma C_{B,2}^\gamma + k_{c,3}^\gamma C_{B,3}^\gamma \\ k_{a,2}^\gamma E_\gamma C_2 \end{bmatrix} - \begin{bmatrix} k_{a,2}^\gamma E_\gamma C_2 \\ k_{b,2}^\gamma C_{B,2}^\gamma + k_{c,2}^\gamma C_{B,2}^\gamma \end{bmatrix} - \frac{1}{M_2} \begin{bmatrix} \frac{1}{Y_2} \\ 0 \end{bmatrix} r_{X,2} V_2 X \quad (5.36)$$

$$\frac{d}{dt} \begin{bmatrix} C_3 \\ C_{B,3}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,3}^\gamma C_{B,3}^\gamma + k_{c,4}^\gamma C_{B,4}^\gamma \\ k_{a,3}^\gamma E_\gamma C_3 \end{bmatrix} - \begin{bmatrix} k_{a,3}^\gamma E_\gamma C_3 \\ k_{b,3}^\gamma C_{B,3}^\gamma + k_{c,3}^\gamma C_{B,3}^\gamma \end{bmatrix} - \frac{1}{M_3} \begin{bmatrix} \frac{1}{Y_3} \\ 0 \end{bmatrix} r_{X,3} V_3 X \quad (5.37)$$

$$\frac{d}{dt} \begin{bmatrix} C_i \\ C_{B,i}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,i}^\gamma C_{B,i}^\gamma + k_{c,i+1}^\gamma C_{B,i+1}^\gamma \\ k_{a,i}^\gamma E_\gamma C_i \end{bmatrix} - \begin{bmatrix} k_{a,i}^\gamma E_\gamma C_i \\ k_{b,i}^\gamma C_{B,i}^\gamma + k_{c,i}^\gamma C_{B,i}^\gamma \end{bmatrix}; \quad i = 4 \text{ to } p \quad (5.38)$$

$$\frac{d}{dt} \begin{bmatrix} C_{p+1} \\ C_{B,p+1}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,p+1}^\gamma C_{B,p+1}^\gamma + n_{p+1,p+2}^\gamma k_{c,p+2}^\gamma C_{B,p+2}^\gamma \\ k_{a,p+1}^\gamma E_\gamma C_{p+1} \end{bmatrix} - \begin{bmatrix} k_{a,p+1}^\gamma E_\gamma C_{p+1} \\ k_{b,p+1}^\gamma C_{B,p+1}^\gamma + k_{c,p+1}^\gamma C_{B,p+1}^\gamma \end{bmatrix} \quad (5.39)$$

$$\frac{d}{dt} \begin{bmatrix} C_i \\ C_{B,i}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,i}^\gamma C_{B,i}^\gamma + \sum_{j=i}^{i+1} n_{ij}^\gamma k_{c,j}^\gamma C_{B,j}^\gamma \\ k_{a,i}^\gamma E_\gamma C_i \end{bmatrix} - \begin{bmatrix} k_{a,i}^\gamma E_\gamma C_i \\ k_{b,i}^\gamma C_{B,i}^\gamma + k_{c,i}^\gamma C_{B,i}^\gamma \end{bmatrix}; \quad i = p+2 \text{ to } p+q-1 \quad (5.40)$$

$$\frac{d}{dt} \begin{bmatrix} C_{p+q} \\ C_{B,p+q}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,p+q}^\gamma C_{B,p+q}^\gamma + n_{p+q,p+q}^\gamma k_{c,p+q}^\gamma C_{B,p+q}^\gamma \\ k_{a,p+q}^\gamma E_\gamma C_{p+q} \end{bmatrix} - \begin{bmatrix} k_{a,p+q}^\gamma E_\gamma C_{p+q} \\ k_{b,p+q}^\gamma C_{B,p+q}^\gamma + k_{c,p+q}^\gamma C_{B,p+q}^\gamma \end{bmatrix} \quad (5.41)$$

In the above, it was assumed that bounded sugars cannot be assimilated by yeast. Therefore, the withdrawal of nutrients was only enforced on the free sugars. If future discoveries suggest otherwise, it is straight forward to add more consumption terms to

the bounded sugars. The form of Eqs. (5.35) - (5.41) is a result of the FP approximation for chain-end scission developed in Section 4.1 where n_{ij}^γ is reproduced here as:

$$n_{ij}^\gamma = \int_{x_i}^{x_{i+1}} \left[\frac{x_{i+1} - v}{x_{i+1} - x_i} \right] \delta(v - [x_j - v_m]) dv + \int_{x_{i-1}}^{x_i} \left[\frac{v - x_{i-1}}{x_i - x_{i-1}} \right] \delta(v - [x_j - v_m]) dv \quad (5.42)$$

Recalling briefly, here v is the continuous DP, x is the pivot DP, and $v_m = 1$ is the DP of the monomer.

Given the necessary initial conditions and parameter values, Eqs. (5.26) - (5.42) constitute an example system of the interlinked population balance and cybernetic modelling framework. A critical component of the model system is the kinetic parameters, particularly those for the hydrolysis of starch as $k_{a,i}^\gamma$, $k_{b,i}^\gamma$ and $k_{c,i}^\gamma$ are known to be DP-dependent, which makes the individual attainment of their values impractical. As far as PBM for enzymatic scission is concerned, the common power law expressions had been used to correlate the rate constants with the DP (Griggs et al., 2012a, 2012b; Hosseini & Shah, 2011a, 2011b). These are simplistic kernels which do not relate to the biochemistry of enzymatic hydrolysis. To model the kinetics of hydrolysis by glucoamylase, the subsite theory (Hiromi, 1970; Hiromi et al., 1973) was adapted in this work (cf. Section G.1 of Appendix G for a summary of the theory). According to this theory, the active site of glucoamylase consists of a tandem array of subsites, each capable of binding to a glucose residue. Although different origins of glucoamylase may contain slightly different number of subsites, the generally encountered value is seven (Fagerström, 1991). This value was adopted here because the subsite of glucoamylase from *S. cerevisiae* (var. *diastaticus*) had not been mapped in the literature. As larger polymers possess higher propensities to collide with glucoamylase molecules, it was assumed that $k_{a,i}^\gamma$ [L/mol/h] increases linearly with the size of the substrate:

$$k_{a,i}^\gamma = \hat{k}_a^\gamma x_i; \quad i = 1 \text{ to } p + q \quad (5.43)$$

where \hat{k}_a^γ [L mol⁻¹ h⁻¹ DP⁻¹] is a constant. With the assumption of rapid equilibrium for the formation of enzyme complexes, the M-M parameter $K_{m,i}^\gamma = k_{b,i}^\gamma / k_{a,i}^\gamma$. Thus, $k_{b,i}^\gamma$ [1/h] ($i = 1$ to $p+q$) can be determined from $K_{m,i}^\gamma$ calculated from the subsite theory:

$$k_{b,i}^\gamma = K_{m,i}^\gamma k_{a,i}^\gamma; \quad K_{m,i}^\gamma = \left[\sum_{J=1}^{N_m^\gamma} K_{J,i}^\gamma \right]^{-1}; \quad K_{J,i}^\gamma = 0.0175 \exp \left(\sum_n^{\text{occ.}} \frac{A_n^\gamma}{RT} \right)_{J,i} \quad (5.44)$$

Here, $K_{J,i}^\gamma$ [mol/L] is the association constants of the i -mer substrate in a binding mode J (which may either be productive or non-productive) and N_m^γ is the total number of subsites of the glucoamylase. In addition, A_n^γ [kcal/mol] is the subsite affinity of the n -th subsite of glucoamylase, expressed in free energy units, $\sum_n^{\text{occ.}}$ implies that the sum is taken for the occupied subsites, and R [kcal/K/mol] as well as T [K] are the universal gas constant and absolute temperature respectively. Using the original subsite theory, $k_{c,i}^\gamma$ [1/h] increases linearly at the lower DP region and reaches a plateau towards the higher DP region. As this trend had thus far only been validated for short chain oligomers, an empirical expression that retains the approximate shape of the original expression for $k_{c,i}^\gamma$ [1/h] from the theory was used to increase the fitting capability:

$$k_{c,i}^\gamma = \hat{k}_c^\gamma \left(\frac{x_i}{\hat{K}_c^\gamma + x_i} \right) + \delta_\gamma; \quad i = 2 \text{ to } p+q; \quad k_{a,i}^\gamma > k_{b,i}^\gamma \gg k_{c,i}^\gamma \quad (5.45)$$

where \hat{k}_c^γ [1/h], \hat{K}_c^γ [DP] and δ_γ [1/h] are constants.

5.2.2 Parameter Identification and Initial Conditions

Identification of the parameters was done in several ways, as summarized in Table 5.2. As only the order of magnitude of $K_{m,i}^\gamma$ matters, in the absence of the subsite map for S .

cerevisiae (var. *diastaticus*) at 303 K (Nakamura et al., 1997), the values of the subsite affinities reported by Hiromi et al. (1973) for glucoamylase produced by *Rhizopus delemar* at 298K were used, cf. Section G.3 of Appendix G. Preliminary studies also indicated that the model outcome was relatively insensitive to the exact values of $k_{a,i}^\gamma$ and $k_{b,i}^\gamma$ provided that the constraint $k_{a,i}^\gamma > k_{b,i}^\gamma \gg k_{c,i}^\gamma$ is fulfilled. Among the parameters tabulated, 18 of them were determined by calibration using the GA (cf. Section 3.4.2). The population size of the algorithm was chosen as five times the number of parameters (Cox, 2005) while the remaining settings were kept as default. Calibration was simultaneously done for five variables, namely the biomass, starch, glucose, ethanol, and glucoamylase concentrations. Here, starch was assumed to be polymers larger than the size of dextrin, i.e. starch was taken to be the total mass of all polymers with DP > 40 (Kearsley & Dziedzic, 1995). The objective function for minimization (J_{opt}) with GA was formulated as shown in Eq. (3.12) of Section 3.4.2. Here, i = biomass, glucose, starch, ethanol and glucoamylase. The objective function was formulated to weigh all quantities of different scales and different number of points equally by setting $\bar{W}_i = 1$. Integration of the ODEs was done using the ‘ode15s’ sub-routine using a simulation time span (= 200 h) which bracketed the experimental data. The optimization was performed on the workstation C (cf. Section 3.1). The time required on average for one complete run is approximately 27 h and the final J_{opt} was 0.1113.

Table 5.2: Values of model parameters used in case study I where the specified initial ranges for calibration using the Genetic Algorithm (GA) were deduced by bracketing the extreme values reported by several similar studies in the literature (Altintas et al., 2002; Gadgil et al., 1996; Jang & Chou, 2013; Kobayashi & Nakamura, 2003, 2004; Ochoa et al., 2007). Calibration was done using the data reported by Nakamura et al. (1997).

Parameter	Unit	Value	Remark
$\mu_{\max,1}$	h^{-1}	0.512	Calibrated using GA within [0.1, 1]
$\mu_{\max,2}$	h^{-1}	$0.88\mu_{\max,1}$	Calculated from $\mu_{\max,1}$ and $\mu_{\max,2}$ reported by Y.-S. Lee, Lee, Chang, and Chang (1995)
$\mu_{\max,3}$	h^{-1}	$0.88\mu_{\max,2}$	Following after the identification of $\mu_{\max,2}$
μ_{γ}	h^{-1}	0.336	Calibrated using GA within [0.1, 1]
K_1	$g\ L^{-1}$	0.200	Calibrated using GA within [0.001, 0.2]
K_2	$g\ L^{-1}$	$3K_1$	Calculated from K_1 and K_2 reported by Y.-S. Lee et al. (1995)
K_3	$g\ L^{-1}$	$3K_2$	Following after the identification of K_2
K_{γ}	$g\ L^{-1}$	3.00	Calibrated using GA within [0.1, 3]
K_{EtOH}	$g\ L^{-1}$	1.06	Calibrated using GA within [1, 15]
K_I^{γ}	$g\ L^{-1}$	0.293	Calibrated using GA within [0.1, 0.9]
Y_1	$g\text{-DW}\ (g\text{-DP1})^{-1}$	0.431	Calibrated using GA within [0.1, 0.5]
Y_2	$g\text{-DW}\ (g\text{-DP2})^{-1}$	$0.95Y_1$	Calculated from Y_1 and Y_2 reported by Y.-S. Lee et al. (1995)
Y_3	$g\text{-DW}\ (g\text{-DP3})^{-1}$	$0.95Y_2$	Following after the identification of Y_2
Y_1^{EtOH}	$g\text{-EtOH}\ (g\text{-DP1})^{-1}$	0.426	Calibrated using GA within [0.1, 0.5]
Y_2^{EtOH}	$g\text{-EtOH}\ (g\text{-DP2})^{-1}$	$1.05Y_1^{EtOH}$	Calculated from Y_1^{EtOH} and Y_2^{EtOH} reported by Y.-S. Lee et al. (1995)

Table 5.2 continued

Parameter	Unit	Value	Remark
Y_3^{EtOH}	g-EtOH (g-DP3) ⁻¹	$1.05Y_2^{EtOH}$	Following after the identification of Y_2^{EtOH}
$k_{e,1}$	h ⁻¹	9.22×10^{-2}	Calibrated using GA within [0.01, 0.2]
$k_{e,2}$	h ⁻¹	$k_{e,1}$	Assumed equal to $k_{e,1}$ for simplicity
$k_{e,3}$	h ⁻¹	$k_{e,1}$	Assumed equal to $k_{e,1}$ for simplicity
$K_{e,1}$	g L ⁻¹	0.931	Calibrated using GA within [0.1, 1]
$K_{e,2}$	g L ⁻¹	$3K_{e,1}$	Calculated from $K_{e,1}$ and $K_{e,2}$ reported by Y.-S. Lee et al. (1995)
$K_{e,3}$	g L ⁻¹	$3K_{e,2}$	Following after the identification of $K_{e,2}$
$\beta_{e,1}$	h ⁻¹	1.24×10^{-2}	Calibrated using GA within [1×10^{-4} , 5×10^{-2}]
$\beta_{e,2}$	h ⁻¹	$\beta_{e,1}$	Following Kompala and Ramkrishna (1986)
$\beta_{e,3}$	h ⁻¹	$\beta_{e,1}$	Following Kompala and Ramkrishna (1986)
M_γ	g mol ⁻¹	3.00×10^5	Value reported by Adam, Latorre - García, and Polaina (2004) for <i>S. cerevisiae</i> (var. <i>diastaticus</i>)
$k_{E,\gamma}$	h ⁻¹	0.248	Calibrated using GA within [0.1, 1]
$K_{E,\gamma}$	g L ⁻¹	2.28	Calibrated using GA within [0.1, 3]
$\beta_{E,\gamma}$	h ⁻¹	1.36×10^{-3}	Calibrated using GA within [1×10^{-4} , 5×10^{-3}]
β_{EtOH}	L ^{η} g ^{-η} h ⁻¹	1.93×10^{-4}	Calibrated using GA within [1×10^{-5} , 2×10^{-4}]
η	-	2.00	Value reported by Kobayashi and Nakamura (2003)

Table 5.2 continued

Parameter	Unit	Value	Remark
\hat{k}_a^γ	L mol ⁻¹ h ⁻¹ DP ⁻¹	5.00×10 ⁶	Chosen so that $k_{c,i}^\gamma/k_{b,i}^\gamma \ll 1$ within the calibration ranges for \hat{k}_c^γ , \hat{K}_c^γ , and δ_γ
\hat{k}_c^γ	h ⁻¹	724	Calibrated using GA within [1, 1500]
\hat{K}_c^γ	DP	955	Calibrated using GA within [1, 5000]
δ_γ	h ⁻¹	196	Calibrated using GA within [0.1, 200]

Table 5.3 gives the initial conditions of the model. The glucoamylase produced by the recombinant strain reported by Nakamura et al. (1997) was unique in that it could not break the α -1,6-glucoside bond of amylopectin in starch and thus 40% of the initial starch loading of 50 g/L remained un-degraded. As such, cell growth stopped as soon as the 30 g/L of the degradable portion was utilized. Since the experimental data was offset by 20 g/L to minimize confusion, the initial starch loading in the simulation was correspondingly selected as $m_s(0) = 30$ g/L to match. Moreover, in their subsequent work (Kobayashi & Nakamura, 2003) for an identical process, the DP of the starch used was reported as 160. As starch in general does not possess a constant molecular weight, it was assumed that this reported value referred to the number-average DP (\overline{M}_n). In addition, as soluble starch was employed by the authors, a relatively narrow distribution with a polydispersity index ($PD = \overline{M}_w/\overline{M}_n = 1.32$) commonly found in the starch literature (Breuninger et al., 2009) was used to characterize the initial distribution in Table 5.3. Meshing was done according to the guidelines established in Section 4.1 for chain-end scission where the values of the meshing parameters p and q were chosen to be $[p, q] = [51, 149]$. This value of p is the minimum required for a total number of pivots = 200. An example code for the simulation is given in Appendix H.

Table 5.3: Initial conditions used in case study I. For the population balance component, the symbol $m_S(0)$ is the initial mass concentration of starch, \overline{M}_n is the number-average DP, and \overline{M}_w is the weight-average DP.

Variable	Unit	Value	Remark
<u>PBM component ($i = 1$ to $p+q$):</u>			
$C_i(0)$	mol L^{-1}	$\int_{v_i}^{v_{i+1}} c(v, 0) dv$	Initial distribution given by Eqs. (3.1) - (3.3) in Section 3.3.1, where: $m_S(0) = 30$ g/L, $\overline{M}_n = 160$, $\overline{M}_w = 212$
$C_{B,i}^y(0)$	mol L^{-1}	0	No bounded species at time = 0
<u>CM component:</u>			
$X(0)$	g L^{-1}	0.1	Given by Nakamura et al. (1997) and Kobayashi and Nakamura (2003)
$A(0)$	g L^{-1}	0	No ethanol at time = 0
$e_1(0)/e_{\max,1}$	-	0.95	Assumed at a high level because yeast was pre-grown on glucose (Nakamura et al., 1997)
$e_2(0)/e_{\max,2}$	-	0.1	Assumed at a low level because yeast was pre-grown on glucose (Nakamura et al., 1997)
$e_3(0)/e_{\max,3}$	-	0.1	Assumed at a low level because yeast was pre-grown on glucose (Nakamura et al., 1997)
$M_\gamma E_\gamma^0(0)$	g L^{-1}	0.01	Assumed at a low level because yeast was pre-grown on glucose (Nakamura et al., 1997)

5.2.3 Simulation Results

The transients of various important quantities in the fermentation broth for both the calibrated model and the experimental data (Nakamura et al., 1997) are shown in Figure 5.2. A number of observations and subtle points are worth mentioning. As yeast was pre-cultured on glucose, the rate of biomass growth at the initial phase was slow as cellular resources were reallocated to the synthesis of glucoamylase. As time progressed, the rate of biomass growth and ethanol production increased due to the availability of glucose

from the hydrolysis of starch. Remarkably, the concentration of glucoamylase could be modelled explicitly, in contrast to previous similar work which obscured the exact quantity of the extracellular enzyme, adopting only its relative intracellular levels (Altintas et al., 2002; Gadgil et al., 1996; Ochoa et al., 2007). The tailing off of glucoamylase beyond 100 h was most likely due to denaturation. Moreover, unlike Kobayashi and Nakamura (2003) who did not showcase the glucose fit against their experimental data, the glucose profile from the interlinked PCM-CM model agreed fairly well with the experimental data. A dual-peak profile was predicted by the interlinked model. Although this was not clearly reflected by the under-sampled experimental data, the second peak was most likely due to the steep decline in the concentration of starch around 60 h as a result of the accumulation of glucoamylase in the broth. While the experimental concentrations of maltose (DP2) and maltotriose (DP3) in the broth were not reported, with the interlinked PBM-CM model, these could be readily predicted. They peaked only when the bulk of starch had been hydrolyzed, and were about two orders of magnitude lower than that of glucose. This is expected for chain-end scission, as it takes time for the large starch molecules to be trimmed down to dimers and trimers. Beyond 100 h, the SSF process was essentially completed. Considering the challenges in obtaining reliable data from bioreactors, the overall quality of fit of the model output to experimental data is very satisfactory. The relatively poorer fit to the starch profile could be attributed to the absence of resistant starch portion in the formulation of starch (Polakovič & Bryjak, 2004), which if present could delay the action of glucoamylase, causing starch to deplete at a slower rate towards the later phase.

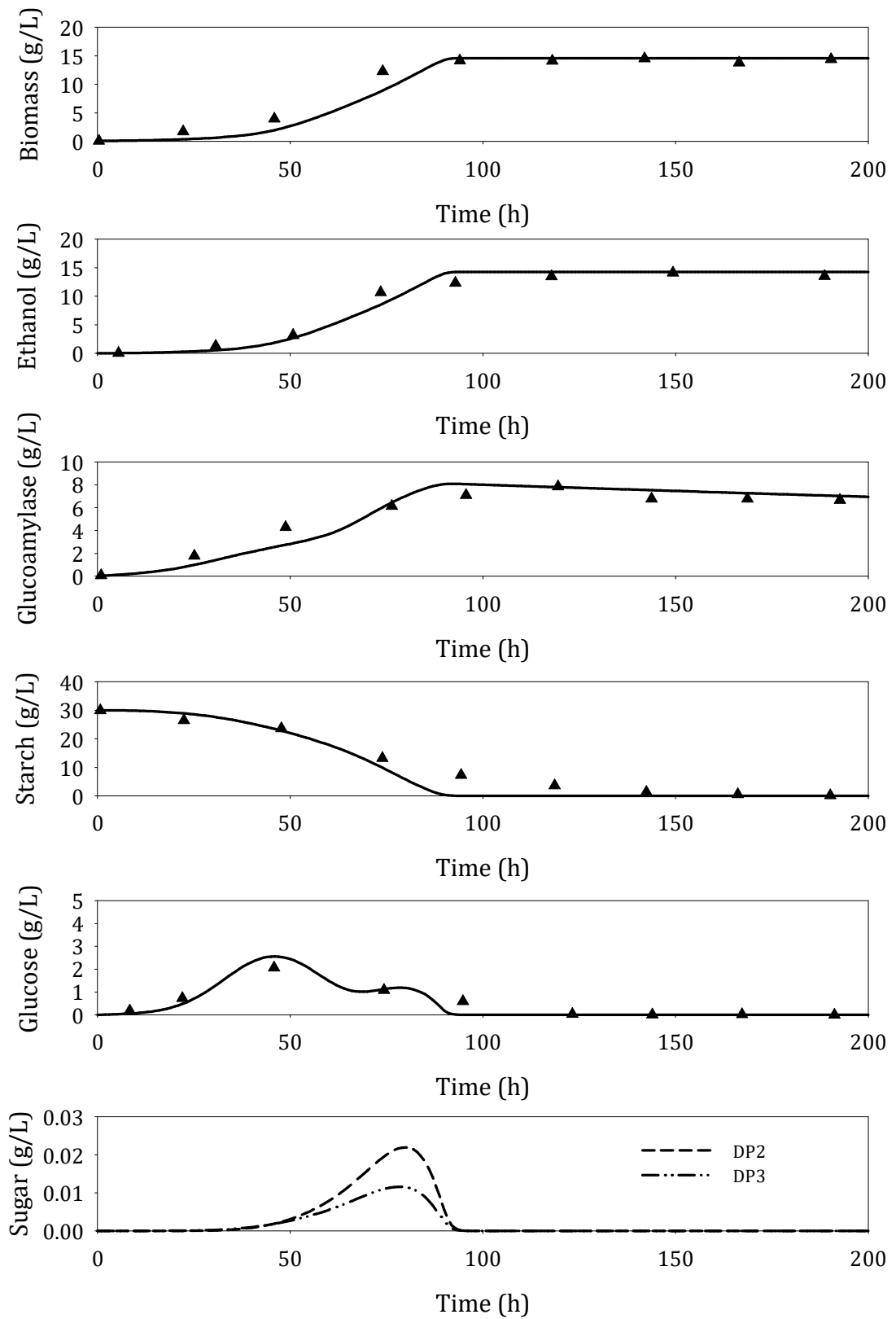


Figure 5.2: Transients of various quantities in the fermentation broth where model predictions are represented by lines and experimental data are represented by symbols. Here, starch (30 g/L) is the sole initial substrate.

Figure 5.3 depicts the temporal evolution of the molar concentration densities of the broth. Such detailed information is a characteristic of the PBM component of the model. From the figure, two key observations could be gleaned. First, the larger the molecular size, the more slowly the molar concentration densities changed with time. This is expected as it would take a longer time for longer polymers to be shortened significantly by cutting one monomer at a time. The second observation is that the absolute magnitudes of the molar concentration densities differed greatly. That for glucose was orders of magnitude larger from the rest despite being utilized concurrently by yeast. This is corroborated by the biomass and glucose profiles in Figure 5.2 before 50 h, which show steeper glucose build-up relative to biomass. Such response can be modified by adjusting the relative abundance of starch to yeast and / or the rate of glucose uptake by the yeast. Seen from another angle, the transient of the number-average DP

$$(\overline{M}_n = \sum_{j=1}^{p+q} x_j [C_j + C_{B,j}^\gamma] / \sum_{j=1}^{p+q} [C_j + C_{B,j}^\gamma])$$

was reduced significantly by the sharp rise in the number of glucose molecules. In contrast, since the removal of glucose units from larger chains did not reduce the molecular weight of the polymer significantly, the

$$\text{weight-average DP } (\overline{M}_w = \sum_{j=1}^{p+q} x_j^2 [C_j + C_{B,j}^\gamma] / \sum_{j=1}^{p+q} x_j [C_j + C_{B,j}^\gamma])$$

displayed a slower decrease over time. Thus, a signature characteristic of chain-scission action on starch is an initial increase in the $PD = \overline{M}_w / \overline{M}_n$ followed by an eventual decrease when the bulk of the starch is converted to glucose. As these polymer properties are associated with the viscosity of the broth which may form a major concern in the design and operation of the bioreactor, results here facilitate its extraction, e.g. through the Mark-Houwink equation (Hiemenz & Lodge, 2007). Certainly, if the starch has a different initial molecular size distribution, the viscosity and the time to reach a set value will be different. Despite this, the qualitative nature of the temporal evolution of the various polysaccharide molecules is expected to be similar.

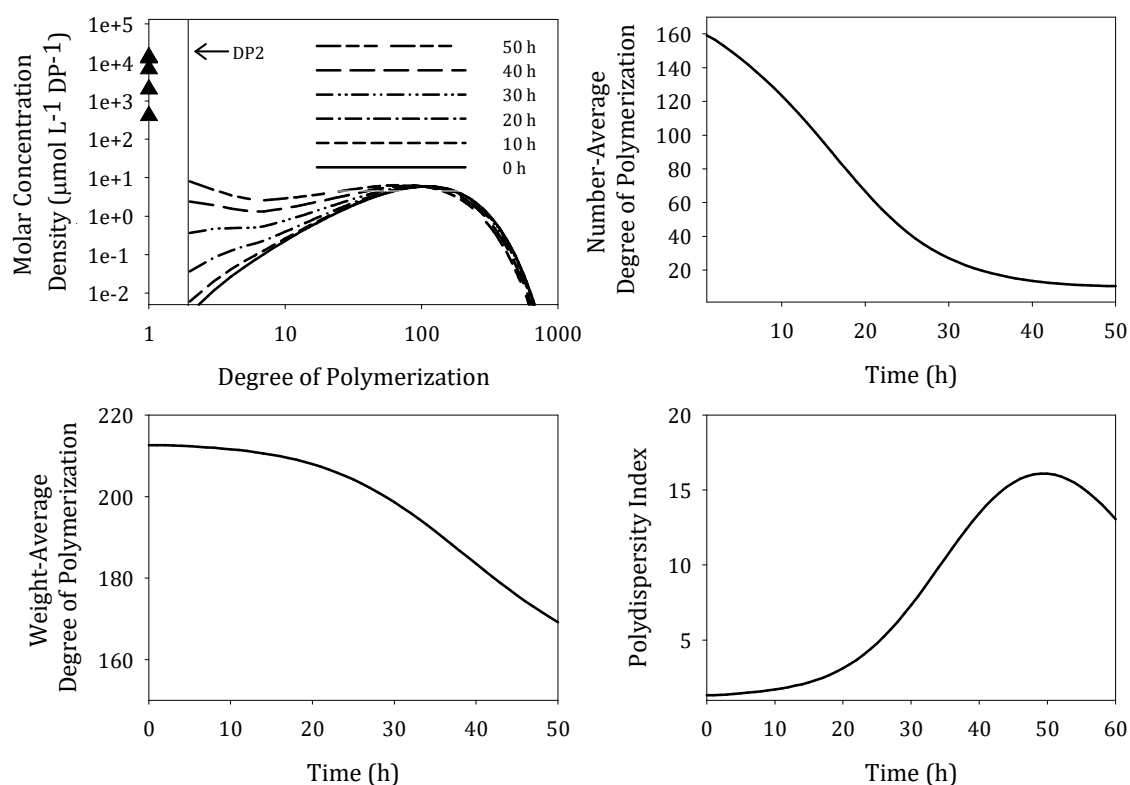


Figure 5.3: Transients of polymer properties corresponding to the system described in Figure 5.2. For the molar concentration density, the triangular symbol represents the concentration density of glucose where the lowest point corresponds to time = 10 h, followed by time = 20 h to time = 50 h in the order of increasing density values.

Using the calibrated model, how yeast would behave was further explored. Although conclusive pronouncement on the model predictions beyond the operating conditions by which the model was calibrated will require further experimental and modelling investigations, the studies below are nonetheless useful for experimental planning. Figure 5.4 shows the model predictions on the effect of adding glucose, or maltose to the starch-containing broth. When glucose was present in addition to starch, preferential utilization of glucose at the onset of fermentation promoted the rapid growth of biomass. For as long as glucose was available in the broth, the synthesis of glucoamylase was repressed and the depletion of starch did not commence. On the other hand, for maltose-starch and maltotriose-starch (not shown) mixtures, the growth of biomass was less vigorous, in part

because the yeast was pre-cultured on glucose. In both cases, the excretion of glucoamylase was not repressed due to the absence of glucose. None of the cases exhibited the classic diauxic growth pattern (Kompala & Ramkrishna, 1986; Monod, 1942) where the biomass would show two distinct exponential growth phases, one for glucose followed by another for the less preferred sugar. This was chiefly due to the continuous supply of glucose by the action of extracellular glucoamylase and the concomitant depletion of the other sugar via scission.

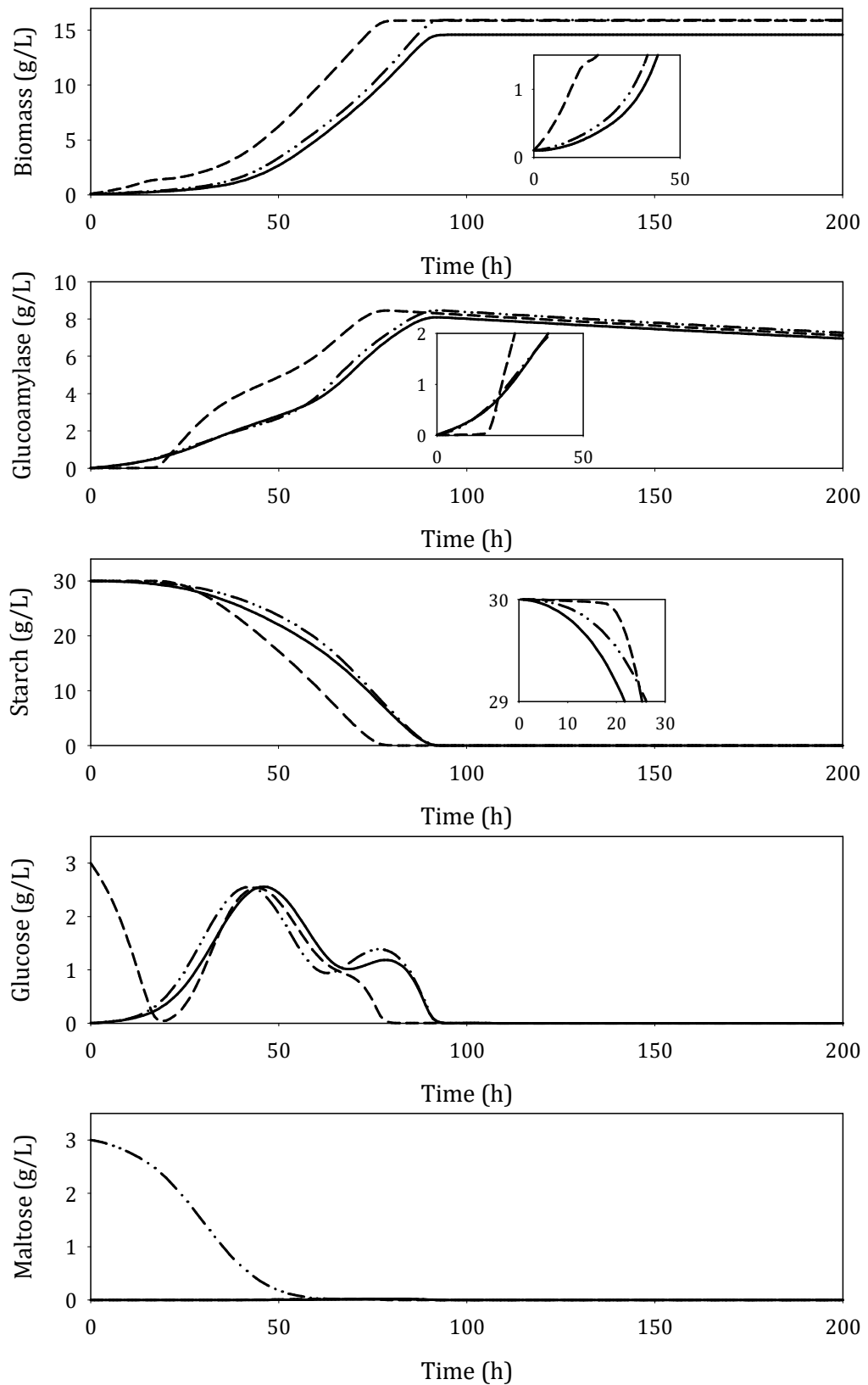


Figure 5.4: Transients of various quantities in the fermentation broth when 3 g of glucose (dashed line), or 3 g of maltose (dashed-double-dotted line) was added in addition to 30 g of starch at time = 0 h. The profile for 30 g starch (solid line) as the sole substrate is given as reference.

Curiously, although the initial presence of glucose in the fermentation broth repressed the synthesis of glucoamylase which is critical for the hydrolysis of starch, Figure 5.4 suggests that overall, the addition of glucose hastened starch hydrolysis. One plausible explanation is that as a result of the initial rapid growth on glucose, more cells were available to produce glucoamylase when the repressive effect of glucose disappeared, resulting in a significant increase in the concentration of glucoamylase within a short time window. This is corroborated by the biomass and the glucoamylase profiles. Such early abundance of glucose was not available when only starch or maltose-starch was present, as it took nearly 50h for the glucose concentrations to peak. Consequently, biomass accumulated more sluggishly, with the corresponding lower secretions of glucoamylase as glucose was being depleted.

The presence of a cybernetic component avails insights beyond the macroscopic parameters above. When maltose and starch are present, the glucose-nursed yeast is forced to shift resources to generate the key enzymes for maltose consumption, while at the same time explore the surrounding for potential glucose availability by secreting some glucoamylase. Such behaviour can be seen in Figure 5.5, which shows a precipitous drop in $e_1/e_{\max,1}$ together with a rise in $e_2/e_{\max,2}$ and $U_{E,\gamma}$, with V_2 kept at the highest level for about 8 h to maximize the assimilation of maltose. Due to the action of glucoamylase on starch, glucose became abundant soon after. Since glucose remains the most favored substrate, the yeast began to redeploy resources to generate enzymes to grow on glucose, as seen in the U-turn in the $e_1/e_{\max,1}$ profile and a rapid rise of $U_{e,1}$ and V_1 , with the concomitant decline in the maltose counterparts. As this switch from maltose to glucose uptake occurred rapidly, the level of $e_2/e_{\max,2}$ only reached a low peak value. Further, as the glucose production rate was much faster than its consumption, the yeast throttled the secretion of glucoamylase, resulting in $U_{E,\gamma}$ showing a maximum. Considering that the

glucoamylase action could not produce any maltotriose (except in the very last stages), the corresponding enzyme for maltotriose utilization was virtually absent and the biomass growth on it completely repressed.

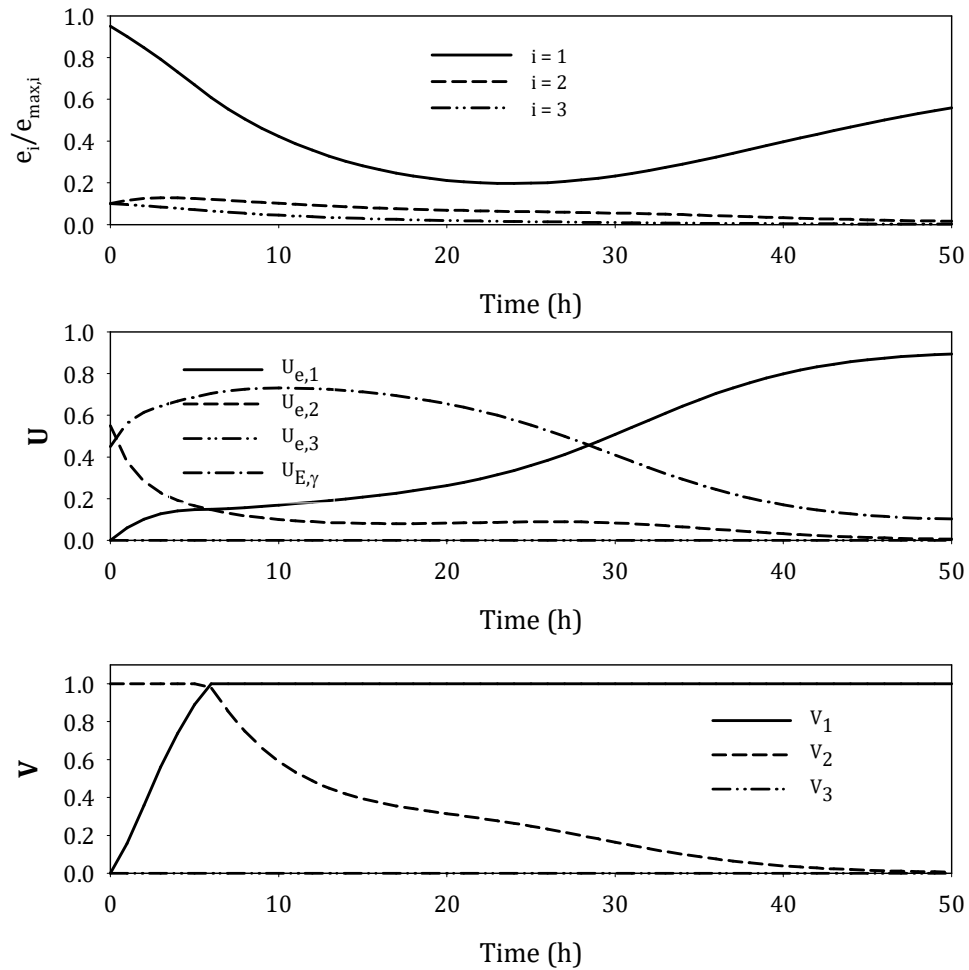


Figure 5.5: Transients of the level of key metabolic enzymes, plus the cybernetic variables U and V corresponding to the maltose-starch mixture presented in Figure 5.4.

When only maltotriose and starch were present initially, a slightly more complicated scenario is expected. Under the action of glucoamylase, maltose would be produced, before being further cut into glucose. Thus, there was a finite period of time when all three consumable simple sugars were present. How would the yeast optimize the choice of the substrate to grow on and the key enzymes to make is not obvious from the outset.

Figure 5.6 shows the various quantities in the fermentation broth when a considerable amount of maltotriose (30 g) was added to 30 g of starch. In general, the ethanol and glucoamylase profiles were similar to the maltose-starch example discussed previously. However, differing from the maltose-starch example, starch was hydrolyzed slowly at the initial phase, an observation which was counter-intuitive as there was no sign of glucoamylase repression in this case. On closer examination, this was due to two reasons. First, the presence of a large amount of maltotriose competed for the availability of glucoamylase, which in turn reduced the amount of free enzymes available to hydrolyze starch. Second, in the chain-end scission of large starch molecules, the resulting larger counterparts are still counted as starch. The ability to capture this slow depletion of starch is a notable feature.

For the remaining sub-figures, as expected, the yeast would initially suppress enzymes for glucose assimilation, produce enzymes to grow on maltotriose, while at the same time secrete glucoamylase to probe the surrounding. When glucose appeared, the yeast quickly switched back to growing on it, as indicated by the rapid rise of V_1 after a mere 10 h, together with $U_{e,1}$ as well as $e_1/e_{\max,1}$ after 30 h. Nevertheless, up to about 60 h, due to the relative abundance of maltotriose compared to glucose, levels of $e_3/e_{\max,3}$, $U_{e,3}$ and V_3 remained noticeable. Interestingly however, within the first 100 h when maltose was present and at times dominant, the corresponding counterparts for maltose were at much lower levels. These observations hint that the optimum strategy of the yeast is to tackle the generally most abundant substrate (maltotriose), grow on the most preferred substrate (glucose) when it becomes available, and ignore the rest. The implication to industrial production is that a significant build-up of a commercially desirable intermediate could be possible if the microbes are kept busy metabolizing other substrates.

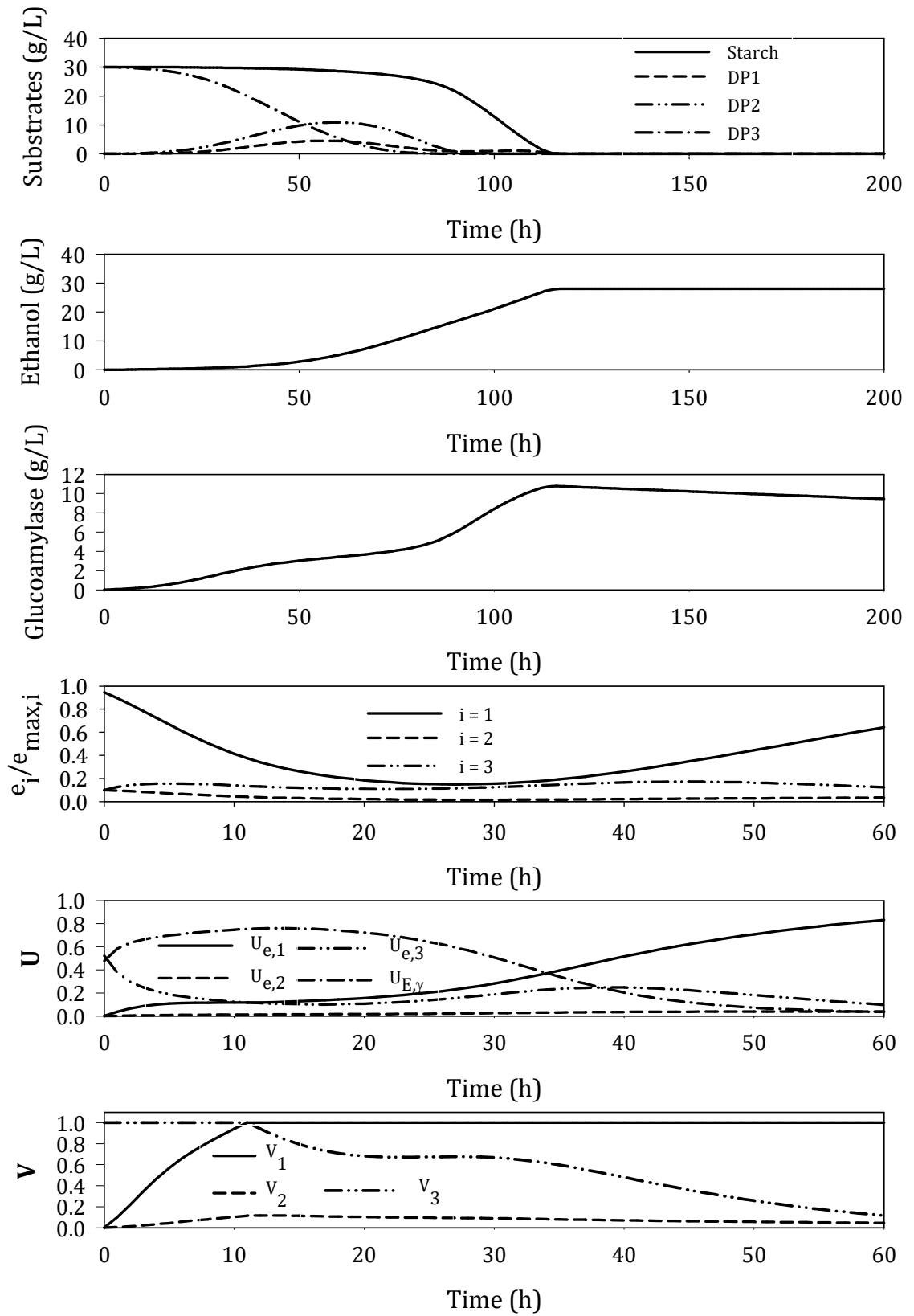


Figure 5.6: Transients of various quantities in the fermentation broth when 30 g of maltotriose was added to 30 g of starch at time = 0 h.

As alluded to previously, the addition of glucose seemed to contribute to an overall increase in the pace of hydrolysis for yeast pre-cultured on glucose. Figure 5.7 examines this further by probing the molar concentration density of the sugars subjected to various initial substrates after 80 h. The figure affirms that the extent of hydrolysis for the glucose-starch initial substrate surpassed that of the other cases. Such an observation was also congruent with the experimental data reported by Altintas et al. (2002) for a similar process. In applications, this efficient breakdown of starch by the addition of glucose to the broth would hasten the abundant supply of nutrients to the fermenting organism, which could promote more timely release of growth-associated products.

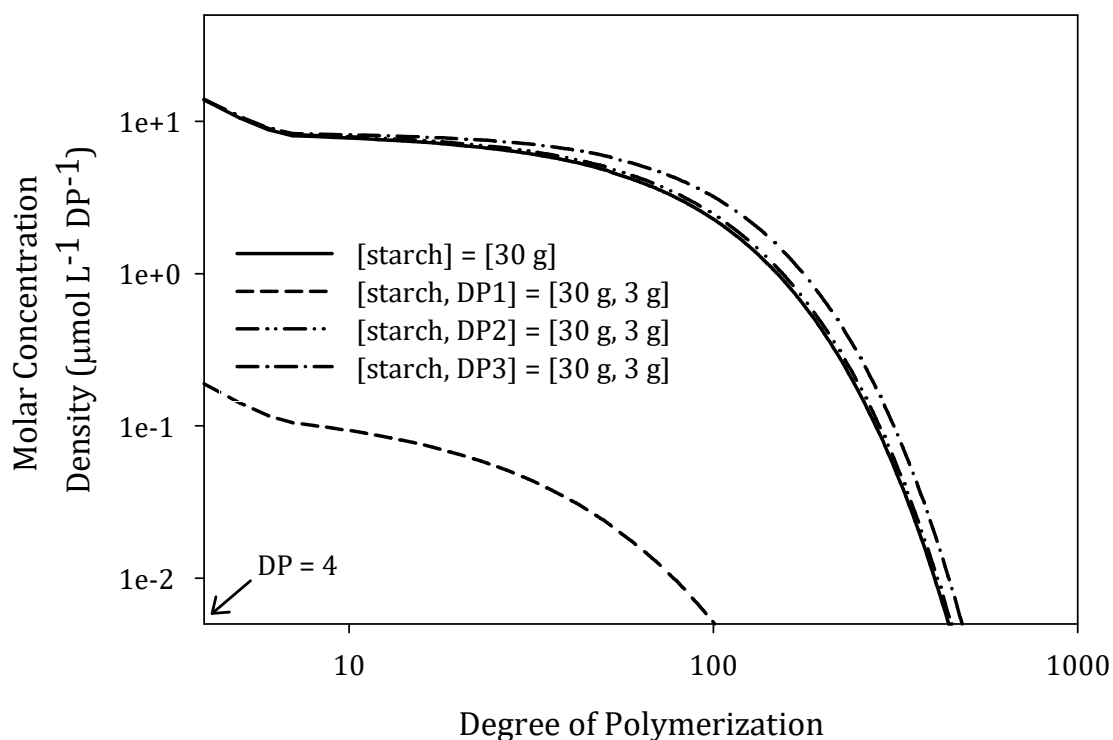


Figure 5.7: Effect of various mixture of initial substrates on the molar concentration density at 80 h.

To cross-check the above implication to product formation, Figure 5.8a shows the effect of adding different mixed substrates to the fermentation broth on ethanol production. As expected, peak ethanol concentration was achieved earlier for the glucose-

containing broth at 78 h, about 14% reduction in the fermentation time compared to the other cases. While adding maltose had a marginal effect in shortening the time to attain maximum ethanol concentration, the addition of maltotriose did not. As such, when yeast is pre-cultured on glucose, to expedite the production of ethanol, it is advantageous for the initial broth to contain glucose. This raises the natural question of whether the conclusion would remain had the yeast been pre-cultured on maltose or maltotriose. Model prediction (details not shown) was affirmative. This can be rationalized as follows: Firstly, if left to its own, a pre-culture broth with maltose or maltotriose will become a glucose broth, as both sugars will induce the excretion of glucoamylase. Secondly, suppose during pre-culture, the glucose level is somehow kept low, and the levels of enzymes for maltose and maltotriose are relatively higher than that for glucose. The introduction of such yeast into a glucose-starch broth will repress enzymes for maltose and maltotriose. Although switching back to glucose-based growth incurs some delay, having glucose readily available bypasses the need to wait for glucoamylase generation and action. Thus, regardless of the pre-culture medium for the yeast, the overall effect is the hastened production of ethanol by adding some glucose to starch. In practical implementations, the starch could be pre-treated before being introduced into the fermentation broth as a substrate. To investigate whether there is an optimum dosage of glucose, Figure 5.8b plots the productivity of ethanol as predicted by the interlinked PBM-CM model against the initial glucose dosage. For fairness, the definition of productivity has been normalized against the initial substrate level and the duration to reach peak ethanol concentration. It turns out that a fairly sharp maximum exists at approximately 7 g of glucose for 30 g of starch. As the effects of the initial substrate quantity had been accounted for, the decline in productivity at a higher glucose dosage is chiefly due to a longer fermentation time, most likely due to product inhibition built into the growth model of the yeast.

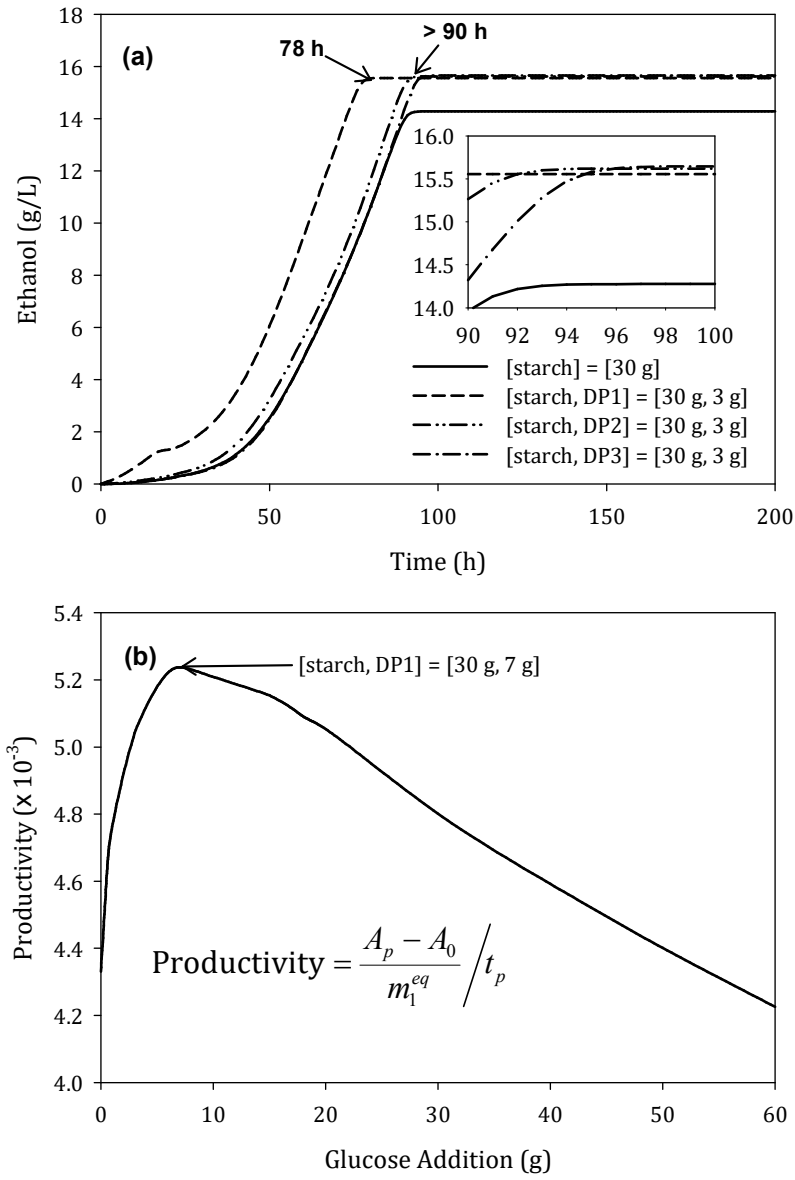


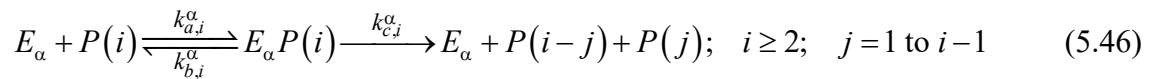
Figure 5.8: Illustrating (a) the ethanol production with various mixtures of initial substrate and (b) the productivity of ethanol as a function of glucose addition, where A is the concentration of ethanol (subscripts ' p ' and ' 0 ' denoting the peak and initial values respectively), m_1^{eq} is the total mass of substrate in the form of glucose equivalent (1 g starch = 1.11 g glucose), and t_p is the time required to achieve peak ethanol concentration.

5.3 Case Study II: Growth of A Glucoamylase and α -amylase Producing Recombinant *S. cerevisiae* on Starch

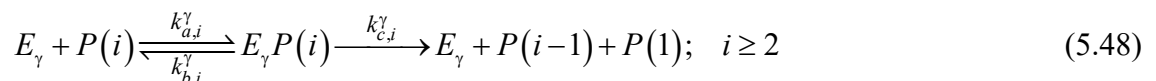
5.3.1 Model Formulation

In the work of De Moraes, Astolfi-Filho, and Oliver (1995), a recombinant yeast, i.e. *S. cerevisiae* YPG/AB, capable of expressing both *Bacillus subtilis* α -amylase and *Aspergillus awamori* glucoamylase as separate poly-peptides was constructed. The performance of this yeast strain in converting starch to alcohol in anaerobic SSF environment was further investigated by Ülgen et al. (2002). To model this SSF system using the interlinked PBM-CM framework, the major differences between the current and the previous case study are: a) the excretion of two different extracellular depolymerases must now be accounted for in the CM component, and consequently b) the two different enzymatic actions on starch must now be reflected in the PBM component, with α -amylase being capable of randomly splitting starch into smaller molecules while glucoamylase releases successive glucose units from the non-reducing ends of starch molecules. Following the previous case study, starch was also modelled here as a linear population of polymers. The hydrolysis mechanisms, involving both α -amylase and glucoamylase, are postulated as:

Hydrolysis by α -amylase:



Hydrolysis by glucoamylase:





where $P(i)$ is the polymer with $DP = i$. Glucoamylase, following the previous case study, does not exhibit single chain attack or multiple attack patterns (Robyt, 2009). Although α -amylases from different origins had been reported to exhibit multiple attack pattern (Kondo, Nakatani, Hiromi, & Matsuno, 1978; Robyt, 2009; Robyt & French, 1967), this was reported by Kondo, Nakatani, Matsuno, and Hiromi (1980) to be negligible for α -amylase originating from *B. subtilis*. Thus, the free glucoamylase and α -amylase are recovered upon successful bond scission. For simplicity, it was assumed that both glucoamylase and α -amylase interact independently with the substrates and that they do not simultaneously bind to the same substrate molecule. As described in Section 5.2.1, the inhibition of enzyme action by glucose is implied in Eqs. (5.47) and (5.49). Similarly, inhibition of enzyme action by starch was not considered here for simplicity.

For the purpose of illustration, the CM framework of Kompala and Ramkrishna (1986) as well as the main assumptions about yeast are retained. Briefly, yeast preferentially metabolizes the DP1 – DP3 sugars in this approximate sequence to produce ethanol and that ethanol consumption during low sugar condition by yeast is ignored as observed in the data by Ülgen et al. (2002). Thus, as far as the growth of biomass (X), production of ethanol (A), and the synthesis of intracellular metabolic enzymes (e_i) are concerned, the form of equations in general do not change. However, because cellular death was reported in the work of Ülgen et al. (2002), a term for cellular death is included in the equation for biomass growth. In addition, because two different depolymerases are now involved, the total amount of $DP = i$ at any time would include the free sugar plus two different types of bounded sugar, i.e. those bounded by α -amylase and glucoamylase respectively. The rate of change in these quantities is briefly given as:

$$\frac{dX}{dt} = \mu X - K_d X; \quad \mu = \sum_{i=1}^3 r_{X,i} V_i; \quad r_{X,i} = \frac{\mu_{\max,i} \left[\frac{e_i}{e_{\max,i}} \right] \left\{ M_i \left[C_i + C_{B,i}^\alpha + C_{B,i}^\gamma \right] \right\}}{\left\{ K_i + M_i \left[C_i + C_{B,i}^\alpha + C_{B,i}^\gamma \right] \right\} \left\{ 1 + A/K_{EtOH} \right\}} \quad (5.50)$$

$$\frac{dA}{dt} = \sum_{i=1}^3 \left[\frac{Y_i^{EtOH}}{Y_i} \right] r_{X,i} V_i X \quad (5.51)$$

$$\frac{d}{dt} \left[\frac{e_i}{e_{\max,i}} \right] = r_{e,i} U_{e,i} - \beta_{e,i} \left[\frac{e_i}{e_{\max,i}} \right] - \mu \left[\frac{e_i}{e_{\max,i}} \right]; \quad r_{e,i} = \frac{k_{e,i} \left\{ M_i \left[C_i + C_{B,i}^\alpha + C_{B,i}^\gamma \right] \right\}}{K_{e,i} + M_i \left[C_i + C_{B,i}^\alpha + C_{B,i}^\gamma \right]}; \quad (5.52)$$

$i = 1 \text{ to } 3$

where K_d [h^{-1}] is rate constant for cellular death, and $C_{B,i}^\alpha$ [mol/L] as well as $C_{B,i}^\gamma$ [mol/L] are the molar concentration of substrates i bounded by α -amylase and glucoamylase respectively.

For *S. cerevisiae* YPG/AB, it was observed that the excretion of glucoamylase was promoted by the abundance of starch and repressed by glucose (Ülgen et al., 2002). This is not surprising, as this is the general observation in the literature concerning microbes capable of excreting depolymerases (Kobayashi & Nakamura, 2003, 2004; Koutinas, Wang, Kookos, & Webb, 2003; Nakamura et al., 1997). Similar observations were also reported for the excretion of α -amylase in the literature (Rothstein, Devlin, & Cate, 1986). For the yeast strain considered here, this was corroborated by the observation that when glucose was depleted, the effect of repression by glucose vanished which consequently caused the concentration of α -amylase to increase (Ülgen et al., 2002). Although the yeast obtained the capability to excrete these enzymes through genetic manipulation, the excretion of these depolymerases similarly utilizes cellular resources (Varner & Ramkrishna, 1999a). Therefore, to capture the regulation of these enzymes using a lumped CM, the following formulation is naturally the simplest approach. As the model must be able to handle the situation where dimeric sugars are also capable of inducing the synthesis of the depolymerases, the rate of glucoamylase and α -amylase synthesis

($r_{E,\alpha}$ and $r_{E,\gamma}$ [g/g-DW/h]) is formulated as a function of the mass concentration of DP \geq

2. The rate of change in the molar concentration of both enzymes with their corresponding enzyme balances can then be written as:

$$\frac{dE_{\alpha}^0}{dt} = \frac{1}{M_{\alpha}} r_{E,\alpha} U_{E,\alpha} X - \beta_{E,\alpha} E_{\gamma}^0; \quad r_{E,\alpha} = \frac{k_{E,\alpha} \sum_{j \geq 2}^{p+q} M_j [C_i + C_{B,i}^{\alpha} + C_{B,i}^{\gamma}]}{K_{E,\alpha} + \sum_{j \geq 2}^{p+q} M_j [C_i + C_{B,i}^{\alpha} + C_{B,i}^{\gamma}]} \quad (5.53)$$

$$\frac{dE_{\gamma}^0}{dt} = \frac{1}{M_{\gamma}} r_{E,\gamma} U_{E,\gamma} X - \beta_{E,\gamma} E_{\gamma}^0 - \beta_{EtOH} A^n E_{\gamma}^0; \quad r_{E,\gamma} = \frac{k_{E,\gamma} \sum_{j \geq 2}^{p+q} M_j [C_i + C_{B,i}^{\alpha} + C_{B,i}^{\gamma}]}{K_{E,\gamma} + \sum_{j \geq 2}^{p+q} M_j [C_i + C_{B,i}^{\alpha} + C_{B,i}^{\gamma}]} \quad (5.54)$$

$$E_{\alpha} = E_{\alpha}^0 - \sum_{j=1}^{p+q} C_{B,j}^{\alpha} \quad (5.55)$$

$$E_{\gamma} = E_{\gamma}^0 - \sum_{j=1}^{p+q} C_{B,j}^{\gamma} \quad (5.56)$$

Note that because two enzymes are now involved, for each enzyme a separate enzyme balance must be written. In addition, as it was reported that the accumulation of ethanol exerts an inhibition effect only on the glucoamylase within the broth (but not α -amylase) (Ülgen et al., 2002), only Eq. (5.54) contains the deactivation term by ethanol. For both the synthesis of intra- and extracellular enzymes, the constitutive term is omitted ($\mathbf{p}_e = \mathbf{0}$ and $\mathbf{p}_E = \mathbf{0}$) following Kompala and Ramkrishna (1986).

The cybernetic variables for controlling the synthesis of intra- and extracellular enzymes, used in Eqs. (5.52) - (5.54), are given as Eq. (5.57) and Eq. (5.58) respectively:

$$U_{e,i} = \frac{r_{X,i} X}{r_{\alpha} X + r_{\gamma} X + \sum_{j=1}^3 r_{X,j} X}; \quad i = 1 \text{ to } 3 \quad (5.57)$$

$$U_{E,\alpha} = \frac{r_{\alpha} X}{r_{\alpha} X + r_{\gamma} X + \sum_{j=1}^3 r_{X,j} X}; \quad U_{E,\gamma} = \frac{r_{\gamma} X}{r_{\alpha} X + r_{\gamma} X + \sum_{j=1}^3 r_{X,j} X} \quad (5.58)$$

Here, yeast is assumed to “grow” indirectly on starch by excreting both α -amylase and glucoamylase, having growth rates of $r_\alpha X$ [g L⁻¹/h] and $r_\gamma X$ [g L⁻¹/h] respectively:

$$r_\alpha = \frac{\mu_\alpha M_\alpha \left[E_\alpha^0 / X \right] \sum_{j \geq 2}^{p+q} M_j \left[C_j + C_{B,j}^\alpha + C_{B,j}^\gamma \right]}{\left\{ K_\alpha + \sum_{j \geq 2}^{p+q} M_j \left[C_j + C_{B,j}^\alpha + C_{B,j}^\gamma \right] \right\} \left[1 + A / K_{EtOH} \right] \left\{ 1 + \left(M_1 \left[C_1 + C_{B,1}^\alpha + C_{B,1}^\gamma \right] / K_I^\alpha \right) \right\}} \quad (5.59)$$

$$r_\gamma = \frac{\mu_\gamma M_\gamma \left[E_\gamma^0 / X \right] \sum_{j \geq 2}^{p+q} M_j \left[C_j + C_{B,j}^\alpha + C_{B,j}^\gamma \right]}{\left\{ K_\gamma + \sum_{j \geq 2}^{p+q} M_j \left[C_j + C_{B,j}^\alpha + C_{B,j}^\gamma \right] \right\} \left[1 + A / K_{EtOH} \right] \left\{ 1 + \left(M_1 \left[C_1 + C_{B,1}^\alpha + C_{B,1}^\gamma \right] / K_I^\gamma \right) \right\}} \quad (5.60)$$

The form of the rate r_α and r_γ are as deliberated in Section F.2 of Appendix F. Here, repression by glucose is achieved by having the inhibition terms of K_I^α [g/L] and K_I^γ [g/L]. As for the cybernetic variable which controls the activities of the enzymes (used in Eqs. (5.50) - (5.51)), they remain the same as shown in the previous case study because the activity of extracellular enzymes is assumed to be unregulated:

$$V_i = \frac{r_{X,i} X}{\max(r_{X,1} X, r_{X,2} X, r_{X,3} X)}; \quad i = 1 \text{ to } 3 \quad (5.61)$$

Establishing the above, and referring to the hydrolysis mechanisms given in Eqs. (5.46) - (5.49), the material balance equations for glucose, maltose, maltotriose and those represented by $i = 4$ to $p+q$ are given as:

$$\frac{d}{dt} \begin{bmatrix} C_1 \\ C_{B,1}^\alpha \\ C_{B,1}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,1}^\alpha C_{B,1}^\alpha + \sum_{j=2}^{p+q} \left(\frac{2}{x_j - 1} \right) k_{c,j}^\alpha C_{B,j}^\alpha + k_{b,1}^\gamma C_{B,1}^\gamma + 2k_{c,2}^\gamma C_{B,2}^\gamma + \sum_{j=3}^{p+q} k_{c,j}^\gamma C_{B,j}^\gamma \\ k_{a,1}^\alpha E_\alpha C_1 \\ k_{a,1}^\gamma E_\gamma C_1 \end{bmatrix} \quad (5.62)$$

$$- \begin{bmatrix} k_{a,1}^\alpha E_\alpha C_1 + k_{a,1}^\gamma E_\gamma C_1 \\ k_{b,1}^\alpha C_{B,1}^\alpha \\ k_{b,1}^\gamma C_{B,1}^\gamma \end{bmatrix} - \frac{1}{M_1} \begin{bmatrix} \frac{1}{Y_1} \\ 0 \\ 0 \end{bmatrix} r_{X,1} V_1 X$$

$$\frac{d}{dt} \begin{bmatrix} C_2 \\ C_{B,2}^\alpha \\ C_{B,2}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,2}^\alpha C_{B,2}^\alpha + \sum_{j=3}^{p+q} \left(\frac{2}{x_j - 1} \right) k_{c,j}^\alpha C_{B,j}^\alpha + k_{b,2}^\gamma C_{B,2}^\gamma + k_{c,3}^\gamma C_{B,3}^\gamma \\ k_{a,2}^\alpha E_\alpha C_2 \\ k_{a,2}^\gamma E_\gamma C_2 \end{bmatrix} \quad (5.63)$$

$$- \begin{bmatrix} k_{a,2}^\alpha E_\alpha C_2 + k_{a,2}^\gamma E_\gamma C_2 \\ k_{b,2}^\alpha C_{B,2}^\alpha + k_{c,2}^\alpha C_{B,2}^\alpha \\ k_{b,2}^\gamma C_{B,2}^\gamma + k_{c,2}^\gamma C_{B,2}^\gamma \end{bmatrix} - \frac{1}{M_2} \begin{bmatrix} \frac{1}{Y_2} \\ 0 \\ 0 \end{bmatrix} r_{X,2} V_2 X$$

$$\frac{d}{dt} \begin{bmatrix} C_3 \\ C_{B,3}^\alpha \\ C_{B,3}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,3}^\alpha C_{B,3}^\alpha + \sum_{j=4}^{p+q} \left(\frac{2}{x_j - 1} \right) k_{c,j}^\alpha C_{B,j}^\alpha + k_{b,3}^\gamma C_{B,3}^\gamma + k_{c,4}^\gamma C_{B,4}^\gamma \\ k_{a,3}^\alpha E_\alpha C_3 \\ k_{a,3}^\gamma E_\gamma C_3 \end{bmatrix} \quad (5.64)$$

$$- \begin{bmatrix} k_{a,3}^\alpha E_\alpha C_3 + k_{a,3}^\gamma E_\gamma C_3 \\ k_{b,3}^\alpha C_{B,3}^\alpha + k_{c,3}^\alpha C_{B,3}^\alpha \\ k_{b,3}^\gamma C_{B,3}^\gamma + k_{c,3}^\gamma C_{B,3}^\gamma \end{bmatrix} - \frac{1}{M_3} \begin{bmatrix} \frac{1}{Y_3} \\ 0 \\ 0 \end{bmatrix} r_{X,3} V_3 X$$

$$\frac{d}{dt} \begin{bmatrix} C_i \\ C_{B,i}^\alpha \\ C_{B,i}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,i}^\alpha C_{B,i}^\alpha + \sum_{j=i+1}^{p+q} \left(\frac{2}{x_j - 1} \right) k_{c,j}^\alpha C_{B,j}^\alpha + k_{b,i}^\gamma C_{B,i}^\gamma + k_{c,i+1}^\gamma C_{B,i+1}^\gamma \\ k_{a,i}^\alpha E_\alpha C_i \\ k_{a,i}^\gamma E_\gamma C_i \end{bmatrix} \quad (5.65)$$

$$- \begin{bmatrix} k_{a,i}^\alpha E_\alpha C_i + k_{a,i}^\gamma E_\gamma C_i \\ k_{b,i}^\alpha C_{B,i}^\alpha + k_{c,i}^\alpha C_{B,i}^\alpha \\ k_{b,i}^\gamma C_{B,i}^\gamma + k_{c,i}^\gamma C_{B,i}^\gamma \end{bmatrix}; \quad i = 4 \text{ to } p$$

$$\frac{d}{dt} \begin{bmatrix} C_{p+1} \\ C_{B,p+1}^\alpha \\ C_{B,p+1}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,p+1}^\alpha C_{B,p+1}^\alpha + \sum_{j=p+2}^{p+q} n_{p+1,j}^\alpha k_{c,j}^\alpha C_{B,j}^\alpha + k_{b,p+1}^\gamma C_{B,p+1}^\gamma + n_{p+1,p+2}^\gamma k_{c,p+2}^\gamma C_{B,p+2}^\gamma \\ k_{a,p+1}^\alpha E_\alpha C_{p+1} \\ k_{a,p+1}^\gamma E_\gamma C_{p+1} \end{bmatrix} \quad (5.66)$$

$$- \begin{bmatrix} k_{a,p+1}^\alpha E_\alpha C_{p+1} + k_{a,p+1}^\gamma E_\gamma C_{p+1} \\ k_{b,p+1}^\alpha C_{B,p+1}^\alpha + k_{c,p+1}^\alpha C_{B,p+1}^\alpha \\ k_{b,p+1}^\gamma C_{B,p+1}^\gamma + k_{c,p+1}^\gamma C_{B,p+1}^\gamma \end{bmatrix}$$

$$\frac{d}{dt} \begin{bmatrix} C_i \\ C_{B,i}^\alpha \\ C_{B,i}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,i}^\alpha C_{B,i}^\alpha + \sum_{j=i}^{p+q} n_{ij}^\alpha k_j^\alpha C_{B,j}^\alpha + k_{b,i}^\gamma C_{B,i}^\gamma + \sum_{j=i}^{i+1} n_{ij}^\gamma k_{c,j}^\gamma C_{B,j}^\gamma \\ k_{a,i}^\alpha E_\alpha C_i \\ k_{a,i}^\gamma E_\gamma C_i \end{bmatrix} \quad (5.67)$$

$$- \begin{bmatrix} k_{a,i}^\alpha E_\alpha C_i + k_{a,i}^\gamma E_\gamma C_i \\ k_{b,i}^\alpha C_{B,i}^\alpha + k_{c,i}^\alpha C_{B,i}^\alpha \\ k_{b,i}^\gamma C_{B,i}^\gamma + k_{c,i}^\gamma C_{B,i}^\gamma \end{bmatrix}; \quad i = p+2 \text{ to } p+q-1$$

$$\frac{d}{dt} \begin{bmatrix} C_{p+q} \\ C_{B,p+q}^\alpha \\ C_{B,p+q}^\gamma \end{bmatrix} = \begin{bmatrix} k_{b,p+q}^\alpha C_{B,p+q}^\alpha + n_{p+q,p+q}^\alpha k_{p+q}^\alpha C_{B,p+q}^\alpha + k_{b,p+q}^\gamma C_{B,p+q}^\gamma + n_{p+q,p+q}^\gamma k_{c,p+q}^\gamma C_{B,p+q}^\gamma \\ k_{a,p+q}^\alpha E_\alpha C_{p+q} \\ k_{a,p+q}^\gamma E_\gamma C_{p+q} \end{bmatrix} \quad (5.68)$$

$$- \begin{bmatrix} k_{a,p+q}^\alpha E_\alpha C_{p+q} + k_{a,p+q}^\gamma E_\gamma C_{p+q} \\ k_{b,p+q}^\alpha C_{B,p+q}^\alpha + k_{c,p+q}^\alpha C_{B,p+q}^\alpha \\ k_{b,p+q}^\gamma C_{B,p+q}^\gamma + k_{c,p+q}^\gamma C_{B,p+q}^\gamma \end{bmatrix}$$

In the above, yeast, following the previous case study, was assumed to only metabolize the free sugars. The expressions for n_{ij}^α and n_{ij}^γ used above are given as:

$$n_{ij}^{\alpha} = \int_{x_i}^{x_{i+1}} \left[\frac{x_{i+1} - v}{x_{i+1} - x_i} \right] \left[\frac{2}{x_j - 1} \right] dv + \int_{x_{i-1}}^{x_i} \left[\frac{v - x_{i-1}}{x_i - x_{i-1}} \right] \left[\frac{2}{x_j - 1} \right] dv \quad (5.69)$$

$$n_{ij}^{\gamma} = \int_{x_i}^{x_{i+1}} \left[\frac{x_{i+1} - v}{x_{i+1} - x_i} \right] \delta(v - [x_j - v_m]) dv + \int_{x_{i-1}}^{x_i} \left[\frac{v - x_{i-1}}{x_i - x_{i-1}} \right] \delta(v - [x_j - v_m]) dv \quad (5.70)$$

where for $j = i$, the first integral on the RHS of Eqs. (5.69) - (5.70) is omitted because the scission of polymers with the size of x_i produces only smaller polymers with the size of $\leq x_i$ (S. Kumar & Ramkrishna, 1996a).

Equations (5.50) - (5.70) constitute the model equations necessary to describe the SSF of starch by a strain capable of excreting two different forms of depolymerases. The rate kernels for both α -amylase and glucoamylase were based on the subsite theory, as had been done in case study I (cf. Sections G.1 – G.2 of Appendix G for the subsite theory for glucoamylase and α -amylase). From the literature, α -amylase from *B. subtilis* possesses eight subsites ($N_m^{\alpha} = 8$) in which the catalytic site is situated between the sixth and the seventh subsite (Iwasa et al., 1974). On the other hand, for glucoamylase from *A. awamori*, the total number of subsites was reported to be seven ($N_m^{\gamma} = 7$) (Sierks, 1988).

The kinetic parameters for hydrolysis are given by the following expressions:

For α -amylase:

$$k_{a,i}^{\alpha} = \hat{k}_a^{\alpha} x_i; \quad i = 1 \text{ to } p + q \quad (5.71)$$

$$k_{b,i}^{\alpha} = K_{m,i}^{\alpha} k_{a,i}^{\alpha}; \quad K_{m,i}^{\alpha} = \left[\sum_{J=1}^{N_m^{\alpha}} K_{J,i}^{\alpha} \right]^{-1}; \quad K_{J,i}^{\alpha} = 0.0175 \exp \left(\sum_n^{\text{occ.}} \frac{A_n^{\alpha}}{RT} \right)_{J,i} \quad (5.72)$$

$$k_{c,i}^{\alpha} = \hat{k}_c^{\alpha} \left(\frac{x_i}{\hat{K}_c^{\alpha} + x_i} \right) + \delta_{\alpha}; \quad i = 2 \text{ to } p + q; \quad k_{a,i}^{\alpha} > k_{b,i}^{\alpha} \gg k_{c,i}^{\alpha} \quad (5.73)$$

For glucoamylase:

$$k_{a,i}^{\gamma} = \hat{k}_a^{\gamma} x_i; \quad i = 1 \text{ to } p + q \quad (5.74)$$

$$k_{b,i}^\gamma = K_{m,i}^\gamma k_{a,i}^\gamma; \quad K_{m,i}^\gamma = \left[\sum_{J=1}^{N_m^\gamma} K_{J,i}^\gamma \right]^{-1}; \quad K_{J,i}^\gamma = 0.0175 \exp \left(\sum_n^{\text{occ.}} \frac{A_n^\gamma}{RT} \right)_{J,i} \quad (5.75)$$

$$k_{c,i}^\gamma = \hat{k}_c^\gamma \left(\frac{x_i}{\hat{K}_c^\gamma + x_i} \right) + \delta_\gamma; \quad i = 2 \text{ to } p+q; \quad k_{a,i}^\gamma > k_{b,i}^\gamma \gg k_{c,i}^\gamma \quad (5.76)$$

Rationale for using the forms of expressions above are as given in the previous case study.

5.3.2 Parameter Identification and Initial Conditions

Identification of the parameters was done in several ways, as summarized in Table 5.4. Since *S. cerevisiae* YPG/AB excretes α -amylase from *B. subtilis* and glucoamylase from *A. awamori*, in the absence of the subsite map at 303 K, the subsite affinities for α -amylase from *B. subtilis* at 298 K (Iwasa et al., 1974) and glucoamylase from *A. awamori* at 323 K (Sierks, 1988) were used here, cf. Section G.3 of Appendix G. Among the parameters tabulated, 32 of them were determined by calibration using the GA. The population size of the algorithm was chosen as five times the number of parameters (Cox, 2005) while the remaining settings were kept as default. Calibration was simultaneously done for five variables, namely the biomass, starch, glucose, ethanol, and reducing sugar concentrations. Here, starch was assumed to be polymers larger than the size of dextrin, i.e. starch was taken to be the total mass of all polymers with DP > 40 (Kearsley & Dziedzic, 1995). The concentration of reducing sugar (RS, [g/L]) was calculated by the following expression:

$$\text{RS}(t) = M_1 \sum_{x_j \leq 40} (C_j + C_{B,j}^a + C_{B,j}^\gamma) \quad (5.77)$$

As every polymer chain possesses only one reducing end which could react with the 3,5 - dinitrosalicylic acid (DNS) reagent (Bernfeld, 1955) and that glucose is used for the preparation of the standard curve in the DNS method for reducing sugar assay, essentially every polymer molecule detectable by the DNS method should be counted as one glucose

molecule; hence the multiplication by the molecular weight of glucose. As the size of the optimization problem using the GA is large (i.e. 32 parameters to be calibrated), $[p, q] = [10, 50]$ was used to reduce the computational time incurred for calibration. The integrity of the optimized parameters was examined by refining the mesh and re-evaluating the objective function, as discussed in the next paragraph. The time required on average for one complete run is approximately 54 h. During the calibration stages, larger weight, i.e. $\bar{W}_i > 1$, was imposed on the biomass concentration to improve the overall fit. The final J_{opt} (Eq. (3.12) of Section 3.4.2), calculated at $\bar{W}_i = 1$ is 0.0917.

Table 5.4: Values of model parameters used in case study II where the specified initial ranges for calibration using the Genetic Algorithm (GA) were deduced by bracketing the extreme values reported by several similar studies in the literature (Altintas et al., 2002; Birol, Önsan, Kırdar, & Oliver, 1998; Gadgil et al., 1996; Jang & Chou, 2013; Kobayashi & Nakamura, 2003, 2004; Ochoa et al., 2007). Calibration was done using the data reported by Ülgen et al. (2002).

Parameter	Unit	Value	Remark
$\mu_{\max,1}$	h^{-1}	0.787	Calibrated using GA within [0.1, 1]
$\mu_{\max,2}$	h^{-1}	$0.88\mu_{\max,1}$	Calculated from $\mu_{\max,1}$ and $\mu_{\max,2}$ reported by Y.-S. Lee et al. (1995)
$\mu_{\max,3}$	h^{-1}	$0.88\mu_{\max,2}$	Following after the identification of $\mu_{\max,2}$
μ_{α}	h^{-1}	0.433	Calibrated using GA within [0.1, 1]
μ_{γ}	h^{-1}	0.419	Calibrated using GA within [0.1, 1]
K_1	g L^{-1}	1.22	Calibrated using GA within [0.01, 1.5]
K_2	g L^{-1}	$3K_1$	Calculated from K_1 and K_2 reported by Y.-S. Lee et al. (1995)
K_3	g L^{-1}	$3K_2$	Following after the identification of K_2

Table 5.4 continued

Parameter	Unit	Value	Remark
K_α	g L^{-1}	2.66	Calibrated using GA within [0.001, 15]
K_γ	g L^{-1}	13.9	Calibrated using GA within [0.001, 15]
K_{EtOH}	g L^{-1}	7.00	Calibrated using GA within [1, 15]
K_I^α	g L^{-1}	4.23	Calibrated using GA within [0.01, 15]
K_I^γ	g L^{-1}	14.6	Calibrated using GA within [0.01, 15]
Y_1	g-DW (g-DP1)^{-1}	0.127	Calibrated using GA within [0.01, 0.15]
Y_2	g-DW (g-DP2)^{-1}	$0.95Y_1$	Calculated from Y_1 and Y_2 reported by Y.-S. Lee et al. (1995)
Y_3	g-DW (g-DP3)^{-1}	$0.95Y_2$	Following after the identification of Y_2
Y_1^{EtOH}	$\text{g-EtOH (g-DP1)}^{-1}$	0.430	Calibrated using GA within [0.01, 0.5]
Y_2^{EtOH}	$\text{g-EtOH (g-DP2)}^{-1}$	$1.05Y_1^{EtOH}$	Calculated from Y_1^{EtOH} and Y_2^{EtOH} reported by Y.-S. Lee et al. (1995)
Y_3^{EtOH}	$\text{g-EtOH (g-DP3)}^{-1}$	$1.05Y_2^{EtOH}$	Following after the identification of Y_2^{EtOH}
$k_{e,1}$	h^{-1}	0.800	Calibrated using GA within [0.01, 1]
$k_{e,2}$	h^{-1}	0.703	Calibrated using GA within [0.01, 1]
$k_{e,3}$	h^{-1}	0.619	Calibrated using GA within [0.01, 1]
$K_{e,1}$	g L^{-1}	3.20	Calibrated using GA within [0.01, 15]
$K_{e,2}$	g L^{-1}	$3K_{e,1}$	Calibrated using GA within [0.01, 15]
$K_{e,3}$	g L^{-1}	$3K_{e,2}$	Calibrated using GA within [0.01, 15]
$\beta_{e,1}$	h^{-1}	1.00×10^{-2}	Calibrated using GA within [1×10^{-4} , 5×10^{-2}]
$\beta_{e,2}$	h^{-1}	$\beta_{e,1}$	Following Kompala and Ramkrishna (1986)
$\beta_{e,3}$	h^{-1}	$\beta_{e,1}$	Following Kompala and Ramkrishna (1986)
M_α	g mol^{-1}	47000	Value reported by Detera and Friedberg (1979) for <i>B. subtilis</i> α -amylase

Table 5.4 continued

Parameter	Unit	Value	Remark
M_γ	g mol^{-1}	83870	Value reported by Fierobe, Mirgorodskaya, Frandsen, Roepstorff, and Svensson (1997) for <i>S. cerevisiae</i> expressing <i>A. awamori</i> glucoamylase
$k_{E,\alpha}$	h^{-1}	0.818	Calibrated using GA within [0.1, 1]
$k_{E,\gamma}$	h^{-1}	0.447	Calibrated using GA within [0.1, 1]
$K_{E,\alpha}$	g L^{-1}	1.64	Calibrated using GA within [0.001, 15]
$K_{E,\gamma}$	g L^{-1}	10.4	Calibrated using GA within [0.001, 15]
$\beta_{E,\alpha}$	h^{-1}	0.385	Calibrated using GA within $[1 \times 10^{-4}, 5 \times 10^{-1}]$
$\beta_{E,\gamma}$	h^{-1}	2.80×10^{-2}	Calibrated using GA within $[1 \times 10^{-4}, 5 \times 10^{-1}]$
β_{EtOH}	$\text{L}^\eta \text{g}^{-\eta} \text{h}^{-1}$	2.18×10^{-2}	Calibrated using GA within $[1 \times 10^{-4}, 5 \times 10^{-2}]$
η	-	2.00	Value reported by Kobayashi and Nakamura (2003)
\hat{k}_a^α	$\text{L mol}^{-1} \text{h}^{-1} \text{DP}^{-1}$	5.00×10^6	Chosen so that $k_{c,i}^\alpha / k_{b,i}^\alpha \ll 1$ within the calibration ranges for \hat{k}_c^α , \hat{K}_c^α , and δ_α
\hat{k}_a^γ	$\text{L mol}^{-1} \text{h}^{-1} \text{DP}^{-1}$	5.00×10^6	Chosen so that $k_{c,i}^\gamma / k_{b,i}^\gamma \ll 1$ within the calibration ranges for \hat{k}_c^γ , \hat{K}_c^γ , and δ_γ
\hat{k}_c^α	h^{-1}	644	Calibrated using GA within [1, 1500]
\hat{k}_c^γ	h^{-1}	734	Calibrated using GA within [1, 1500]
\hat{K}_c^α	DP	10472	Calibrated using GA within [1, 20×10^3]
\hat{K}_c^γ	DP	49.8	Calibrated using GA within [1, 5000]
δ_α	h^{-1}	6.42	Calibrated using GA within [0.1, 200]
δ_γ	h^{-1}	18.1	Calibrated using GA within [0.1, 200]
K_d	h^{-1}	1.12×10^{-2}	Calibrated using GA within $[1 \times 10^{-4}, 5 \times 10^{-2}]$

Table 5.5 gives the initial conditions of the model. In the work of Ülgen et al. (2002), soluble starch was used. Since no further particulars about the starch employed was given, the number-average DP, $\overline{M}_n = 160$ and a polydispersity index (PD = $\overline{M}_w / \overline{M}_n = 1.32$) as used in case study I were retained here for simplicity. Moreover, the soluble starch was autoclaved at 121°C for 45 mins, and thus 1.68 g/L of reducing sugar was initially present. Since the initial concentration of glucose was reported to be 0.79 g/L, for simplicity the remaining amount of reducing sugar (0.89 g/L) was assumed to be the DP2 sugar. According to the principle employed in Eq. (5.77), this exact amount of the DP2 sugar was calculated to be 1.69 g/L. Meshing was done according to the guidelines established in Section 4.2 for simultaneous random and chain-end scissions. Upon obtaining the optimized parameters as described above, the model was re-run using a finer mesh, i.e. $[p, q] = [51, 149]$ where the value of p is the minimum required for a total number of pivots = 200. In order to ascertain that the coarser mesh employed during calibration did not affect the integrity of the calibrated parameters, the objective function J_{opt} for the new mesh was examined and found to be generally indifferent to the choice of $[p, q]$ i.e. the new $J_{opt} = 0.0920$ vs. the old $J_{opt} = 0.0917$ (less than 0.5% deviation). An example code for the simulation is given in Appendix I.

Table 5.5: Initial conditions used in case study II. For the population balance component, the symbol $m_S(0)$ is the initial mass concentration of starch, \overline{M}_n is the number-average DP, and \overline{M}_w is the weight-average DP.

Variable	Unit	Value	Remark
<u>PBM component ($i = 1$ to $p+q$):</u>			
$C_1(0)$	mol L ⁻¹	$m_1(0)/M_1$	Initial distribution given by Eqs. (3.1) - (3.3) in Section 3.3.1, where $m_S(0) = 38.4$ g/L, $\overline{M}_n = 160$, $\overline{M}_w = 212$, $m_1(0) = 0.79$ g/L, $m_2(0) = 1.69$ g/L, and $\int_{v_i}^{v_{i+1}} c(v,0)dv \approx 0$ for $i = 1$ to 2.
$C_2(0)$	mol L ⁻¹	$m_2(0)/M_2$	
$C_i(0)$	mol L ⁻¹	$\int_{v_i}^{v_{i+1}} c(v,0)dv$ $i = 3$ to $p+q$	
$C_{B,i}^\alpha(0)$	mol L ⁻¹	0	No bounded species at time = 0
$C_{B,i}^\gamma(0)$	mol L ⁻¹	0	No bounded species at time = 0
<u>CM component:</u>			
$X(0)$	g L ⁻¹	0.1	Given by Ülgen et al. (2002)
$A(0)$	g L ⁻¹	0	No ethanol at time = 0
$e_1(0)/e_{\max,1}$	-	0.9	Assumed at a high level because yeast was pre-grown on glucose (Ülgen et al., 2002)
$e_2(0)/e_{\max,2}$	-	0.1	Assumed at a low level because yeast was pre-grown on glucose (Ülgen et al., 2002)
$e_3(0)/e_{\max,3}$	-	0.1	Assumed at a low level because yeast was pre-grown on glucose (Ülgen et al., 2002)
$M_\alpha E_\alpha^0(0)$	g L ⁻¹	1.42×10^{-1}	Calibrated using GA within a low concentration range [1×10^{-5} , 0.15] because yeast was pre-grown on glucose (Ülgen et al., 2002)
$M_\gamma E_\gamma^0(0)$	g L ⁻¹	4.21×10^{-2}	Calibrated using GA within a low concentration range [1×10^{-5} , 0.15] because yeast was pre-grown on glucose (Ülgen et al., 2002)

5.3.3 Simulation Results

The transients of various important quantities in the fermentation broth for both the calibrated model and the experimental data (Ülgen et al., 2002) are shown in Figure 5.9.

In general, the model predictions agreed well with the experimental data. The typical profiles encountered in fermentation studies for biomass, ethanol, and starch concentrations were predicted by the model. One notable feature of the interlinked PBM-CM framework is the ability to model the transient of the concentration of the reducing sugar systematically. In the past where CM was employed to model SSF (Altintas et al., 2002; Ochoa et al., 2007), reducing sugar was modelled as a lumped entity and the resulting fit was not benchmarked against any form of experimental data. Here, because of the ability to account for the complete polymer distribution, the notion of reducing sugar has a more solid definition (Saqib & Whitney, 2011), and is even amenable for the particular type of experimental procedure employed, c.f. Eq. (5.77) for the definition of reducing sugar obtained through the DNS procedure. For systems involving α -amylase where glucose is not the main product, the ability to model the transient of reducing sugar in the broth is important to observe the dynamics of hydrolysis by α -amylase (Birol et al., 1998). From the figure, although the experimental data for both glucose and reducing sugar are slightly noisy, the general trend of the dynamic profiles was captured satisfactorily.

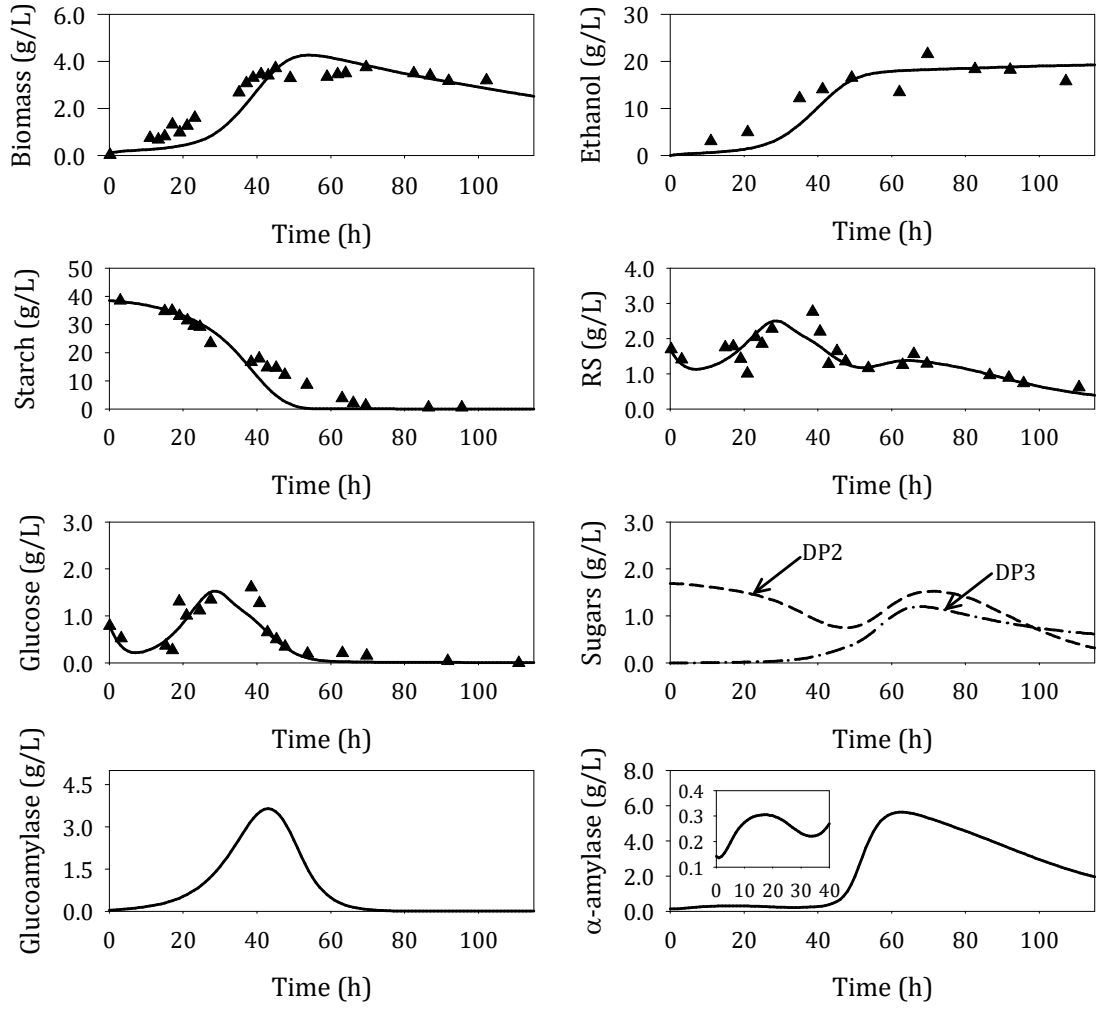


Figure 5.9: Transients of various quantities in the fermentation broth where model predictions are represented by lines and experimental data (Ülgen et al., 2002) are represented by symbols. Here, starch (38.4 g/L), glucose (0.79 g/L) and maltose (1.69 g/L) are initially present. Soluble starch with $\left[\overline{M}_n, \overline{M}_w, N, PD\right] = [160, 212, 878, 1.325]$ was used here.

Although the model was not rigorously calibrated, i.e. only a single set of experimental data as shown in Figure 5.9 was used for calibration, several informative features could nonetheless be gleaned from the predicted transients of the concentrations of α -amylase and glucoamylase. First, α -amylase was produced in small amounts during the first 50 h of the fermentation and remained relatively constant until the stationary phase where its concentration then increased rapidly. In the work of Ülgen et al. (2002), while the authors

mentioned that the measurement of the activity of the depolymerases may be inaccurate due to the compounding effects of both α -amylase and glucoamylase, the general trend displayed for α -amylase was coherent with the prediction in Figure 5.9 in that the activity of α -amylase also increased towards the stationary phase. Furthermore, in a similar work by Birol, Kirdar, and Önsan (2002), this trend for α -amylase excretion was also observed. As a result of this rapid increment in the concentration of α -amylase in the stationary phase, the concentration of reducing sugar exhibited a second peak between 60 to 80 h. During this phase, the consumable sugars which constitute the category of reducing sugar consisted of primarily the DP2 and DP3 sugars as glucose was depleted rapidly by the large amount of biomass present. Second, the concentration of glucoamylase as predicted by the interlinked model increased during the exponential phase of the yeast growth and was deactivated at high ethanol concentration. Such a prediction is again congruent with the general observation encountered in the studies by Ülgen et al. (2002) and Birol et al. (2002). This low/negligible concentration of glucoamylase beyond 60 h also caused glucose to not accumulate during this time period.

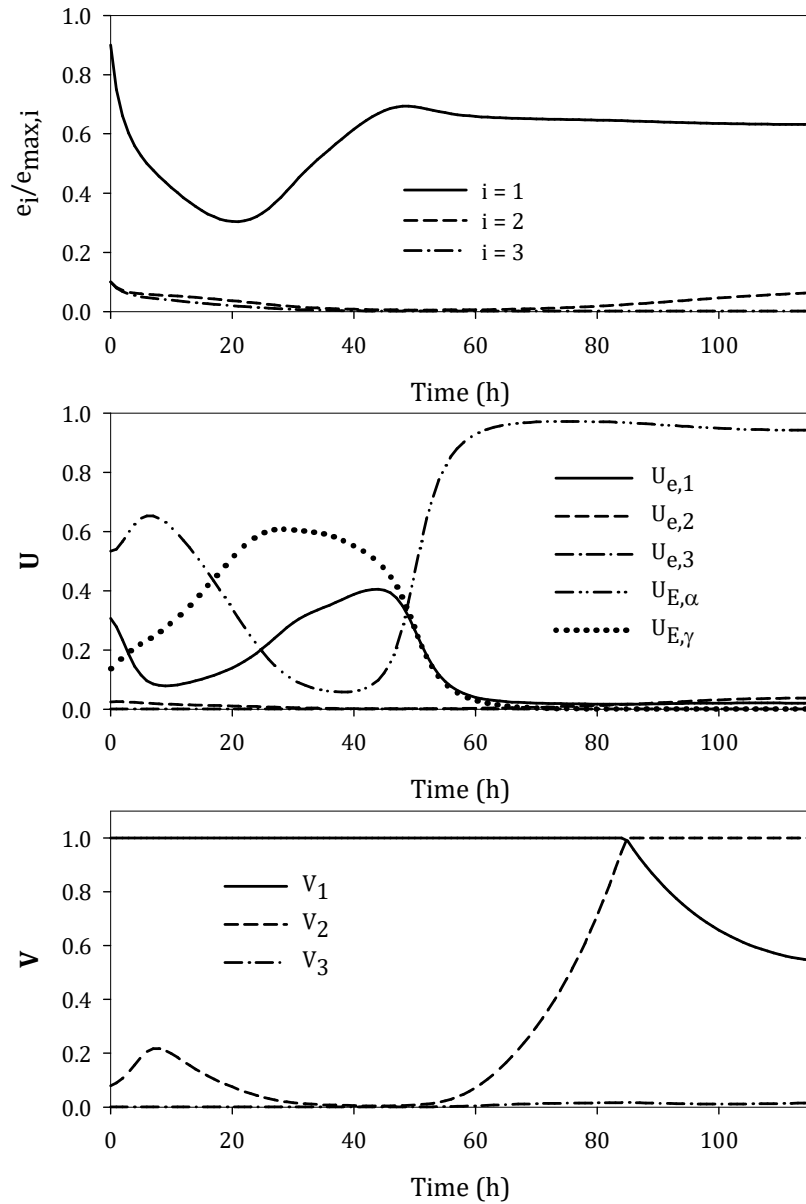


Figure 5.10: Transients of the level of key metabolic enzymes, plus the cybernetic variables U and V corresponding to the simulation presented in Figure 5.9.

Figure 5.10 shows the corresponding level of key metabolic enzymes and the cybernetic variables U and V for the results presented in Figure 5.9. Upon further examination, these figures seem to suggest that yeast manages the excretion of α -amylase and glucoamylase within the fermentation broth according to a particular strategy. At the initial phase, both α -amylase and glucoamylase were excreted by yeast in small amounts where cellular resources during this phase were increased gradually for inducing their

production (see profiles of $U_{E,\alpha}$ and $U_{E,\gamma}$). Although the production of these two enzymes are generally repressed by glucose, the initial amount of glucose present was not sufficiently significant to initiate a complete repression of their production. The role of α -amylase when confronting starch is to break the polymers such that more non-reducing ends are available for glucoamylase attack (Fujii & Kawamura, 1985). As only a small amount of α -amylase is required to liquefy starch, its production was soon after repressed (albeit not completely to sustain its presence in the broth) while yeast focused on producing more glucoamylase which is directly responsible for generating the preferred substrate, i.e. glucose. Because the extent of glucose repression on glucoamylase production is lesser than that for α -amylase (Ülgen et al., 2002), the production of glucoamylase continued despite the rise of glucose level until the concentration of glucose peaked at approximately 30 h. The rise of glucose level also consequently caused the cells to channel more resources for the production of the key enzyme for metabolizing glucose. Soon after the concentration of ethanol reached a significant level beyond 43 h, the concentration of glucoamylase declined rapidly because of the deactivation by ethanol. Deactivation of α -amylase by ethanol does not occur according to Ülgen et al. (2002). The decline in the concentration of glucoamylase by ethanol deactivation was accompanied by a rapid decline in $U_{E,\gamma}$ as cells find it infeasible to continue the induction of glucoamylase synthesis. Under such circumstance where the main enzyme for generating the preferred substrate (glucose) is not available, the next best strategy to generate glucose is by inducing the production of α -amylase. The model prediction validates this point, where a major portion of the cellular resources was allocated for the induction of α -amylase, i.e. $U_{E,\alpha}$ is large. From the profiles for the cybernetic variable V , yeast fully activated the key metabolic enzyme for the consumption of the DP2 sugar beyond 85 h. Even though V_2 is large, growth on the DP2 sugar was still inefficient due to the low level of $e_2/e_{\max,2}$. Because glucose was also produced by α -

amylase (albeit not in very large quantities and thus did not accumulate due to rapid consumption), V_1 was not fully repressed. The foregoing observations seem to suggest that the yeast's strategy for managing the excretion of α -amylase and glucoamylase is to optimize the generation of glucose. Regardless of whether other nutrients (i.e. the DP2 and DP3 sugars) could be used to support growth, the yeast's main priority would be to allocate cellular resources towards this goal. Considering that the model was not rigorously calibrated, the capability of the framework to yield such informative predictions is remarkable.

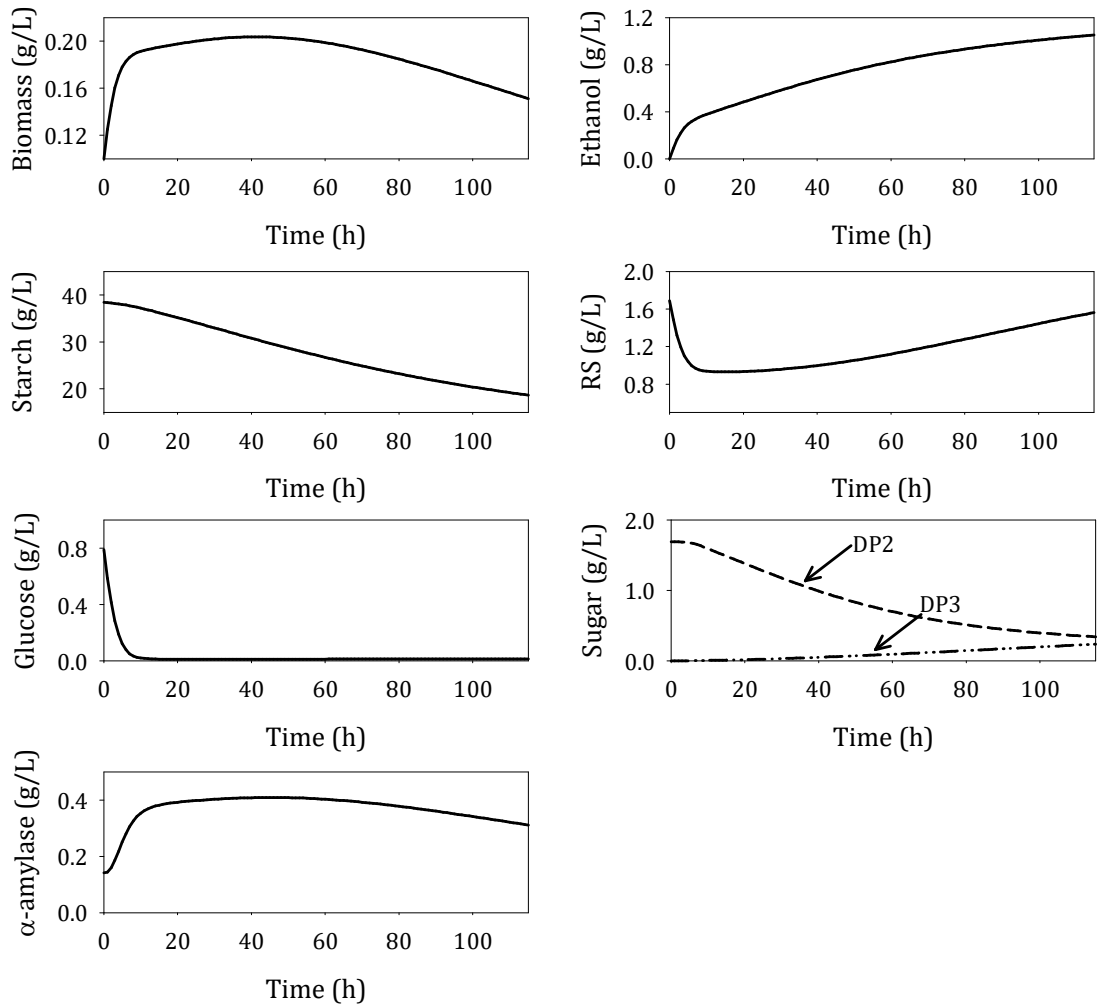


Figure 5.11: Transients of various quantities in the fermentation broth when only α -amylase is produced by the yeast. Soluble starch with $[\overline{M}_n, \overline{M}_w, N, PD] = [160, 212, 878, 1.325]$ was used here.

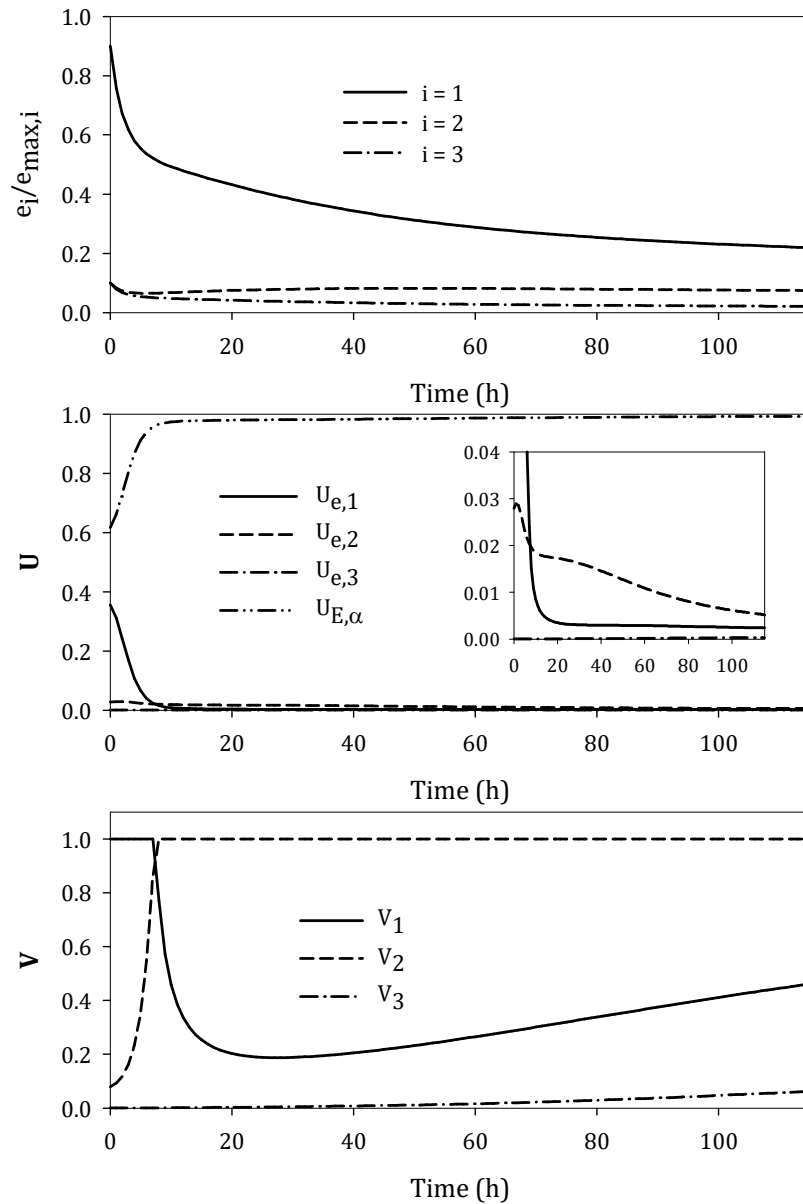


Figure 5.12: Transients of the level of key metabolic enzymes, plus the cybernetic variables U and V corresponding to the simulation presented in Figure 5.11

The results above seem to suggest that both α -amylase and glucoamylase were required for the efficient SSF of starch. To further validate this observation, it is desirable to observe whether having yeast to excrete only either α -amylase or glucoamylase is a sub-optimal strategy for SSF. First, the simulation was re-run by knocking out the ability of the yeast to excrete glucoamylase, i.e. by removing the respective terms in the equations. Thus, in this case, the yeast strain only excretes α -amylase. Figures 5.11 – 5.12

show the transients of the various quantities in the fermentation broth and the level of key metabolic enzymes plus the cybernetic variables U and V . From the figures, it is apparent that without glucoamylase, the generation of glucose was severely retarded. This could be due to two reasons. First, although the hydrolysis of starch by α -amylase also releases glucose, the amount released was immediately consumed by the yeast and thus it did not accumulate in the broth. Second, the amount of α -amylase excreted by the yeast was insufficient to actually produce a significant amount of glucose in the broth. Even though the cells under such a glucose deficient environment decided to allocate almost all of their resources for the production of α -amylase (high $U_{E,\alpha}$), the concentration of α -amylase was not high within the broth because the amount of biomass present was insufficient to promote the release of a larger amount of α -amylase. Such a slow biomass growth was due to the inefficient growth on primarily the DP2 sugar in a glucose deficient environment. In principle, had the yeast been suitably acclimatized, it might have a different set of system parameters that enabled it to grow more efficiently on the DP2 and DP3 sugars. For the current glucose-addicted strain, this was not the case, and the SSF process would be inefficient without glucoamylase producing capabilities.

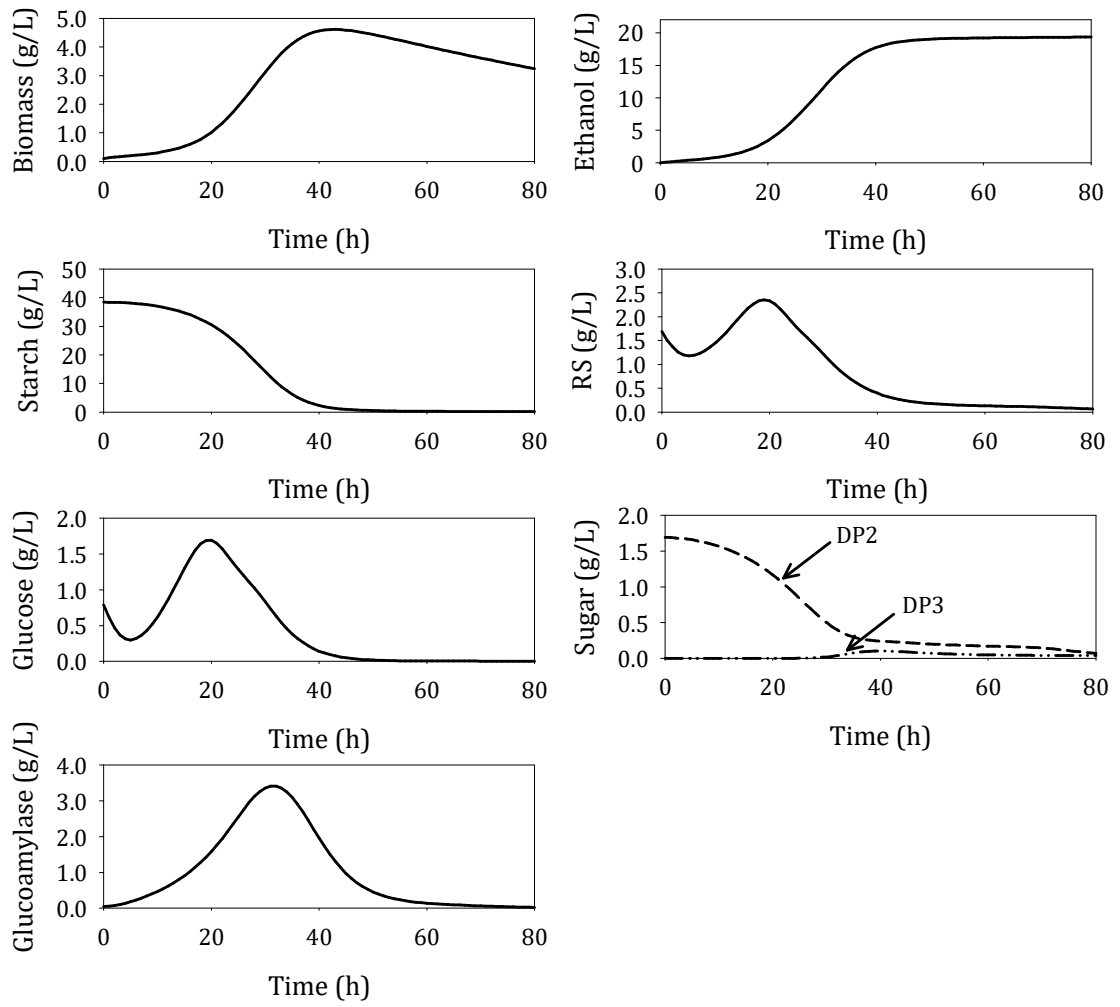


Figure 5.13: Transients of various quantities in the fermentation broth when only glucoamylase is produced by the yeast. Soluble starch with $[\overline{M}_n, \overline{M}_w, N, PD] = [160, 212, 878, 1.325]$ was used here.

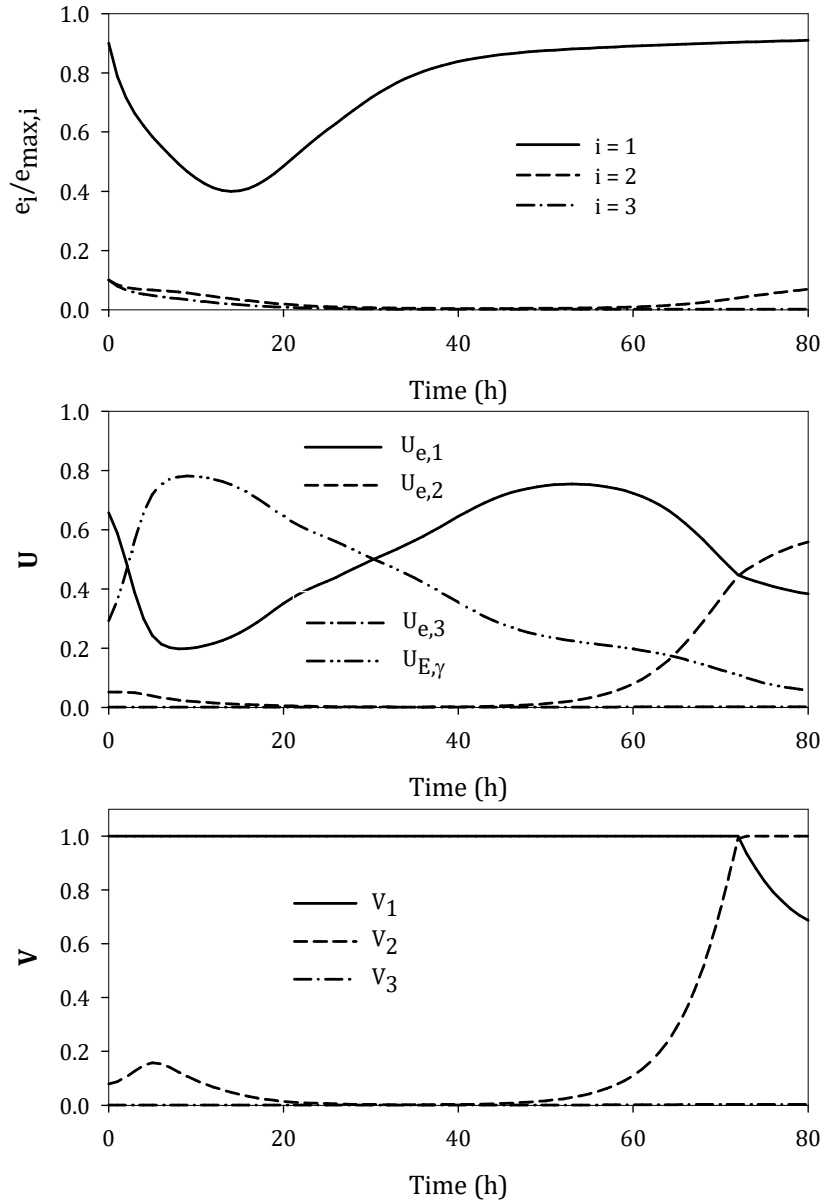


Figure 5.14: Transients of the level of key metabolic enzymes, plus the cybernetic variables U and V corresponding to the simulation presented in Figure 5.13

Figures 5.13 – 5.14 show the model predictions when the yeast excretes only glucoamylase within the broth. On first glance, it may seem that having yeast to solely excrete glucoamylase might be a viable SSF strategy. Because of the action of glucoamylase, the amount of glucose was sufficient during the first 40 h to promote efficient growth of biomass. At the initial period, $U_{E,\gamma}$ increased to induce the production of glucoamylase. When the concentration of glucose in the broth reached a significant

level at approximately 10 h, the synthesis of glucoamylase was repressed and the synthesis of $e_1/e_{\max,1}$ was induced thereafter. After glucose was depleted from the broth beyond 50 h, the cells started to induce the production of the key enzyme for metabolizing the DP2 sugar and began to grow on this sugar as inferred from the increasing value of V_2 . As the growth on the DP2 sugar was inefficient due to the relatively low level of $e_2/e_{\max,2}$, its effect on the fermentation was not noticeable because the SSF is essentially completed during that time period. One point to note here is that similar to the case presented in Figures 5.11 – 5.12, the yeast's strategy for survival is always to excrete the depolymerization enzyme for the generation of consumable substrates. The current results raise the natural question of whether there is a need at all for the excretion of α -amylase, since Figures 5.13 – 5.14 seem to suggest that glucoamylase alone is sufficient for the SSF of starch. The following discussion addresses this issue.

To investigate whether having the yeast to excrete only glucoamylase could possibly be the best strategy for SSF, Figure 5.15 shows the transients of various quantities in the fermentation broth for the SSF of starch with different distribution. Other than that used previously, i.e. $[\overline{M}_n, \overline{M}_w, N, PD] = [160, 212, 878, 1.325]$ (denoted as starch C), the polymer with $[\overline{M}_n, \overline{M}_w, N, PD] = [10000, 10100, 22496, 1.01]$ represents a hypothetical starch with a narrow distribution (referred to as starch A) while the polymer with $[\overline{M}_n, \overline{M}_w, N, PD] = [4100, 5430, 22496, 1.324]$ represents the sweet potato starch (Breuninger et al., 2009) (referred to as starch B here). From the figure, clearly the differences in the distribution of starch affect the performance of the SSF when yeast is capable of only excreting glucoamylase. For starch C, i.e. soluble starch which is starch that had been partially hydrolyzed, SSF is efficient because there are sufficient non-reducing ends for the glucoamylase to act on to produce glucose for efficient growth and

fermentation. However, when starch with a larger N and larger average DPs was used (i.e. starch A or B), the depletion of starch was drastically slower. This could be rationalized because the starch now contains more larger molecules, and in the chain-end scission of large starch molecules, the resulting larger counterparts are still counted as starch. From the figure also, the depletion of starch A was slower than that for starch B. This is because the bulk of the polymer mass for starch A is situated at a higher average ($\overline{M}_w = 10100$) than that for starch B ($\overline{M}_w = 5430$), and therefore glucoamylase took more time to hydrolyze starch A until the molecules were no longer counted as starch. Because of the slower depletion of starches A and B, other quantities in the fermentation broth were also consequently affected. For starches A and B, having more large molecules (within a fixed total mass) reduces the number of non-reducing ends for glucoamylase attack. Hence the generation of glucose was sluggish and consequently the growth of biomass and ethanol production were slow. Moreover, as chain-end scission produces primarily glucose, the concentrations of other sugars in the broth were low and the reducing sugar in this case comprised mainly of glucose. The results here show that having the yeast to excrete glucoamylase alone may not always be adequate because the performance of the SSF in such a case would be affected by the type of starch employed.

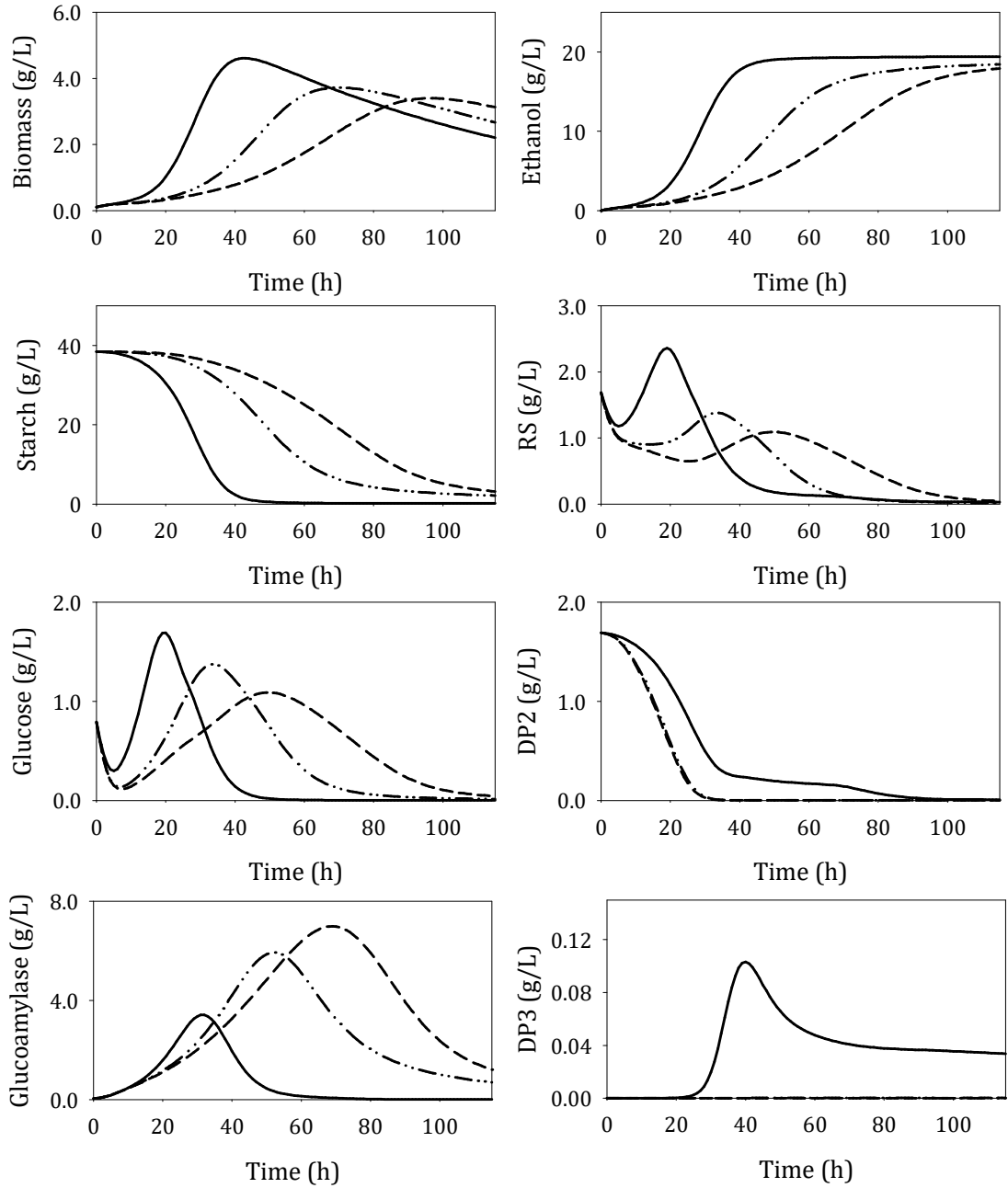


Figure 5.15: Transients of various quantities in the fermentation broth when only glucoamylase is produced by the yeast. Here, dashed lines represent the results for starch A with $[\overline{M}_n, \overline{M}_w, N, PD] = [10000, 10100, 22496, 1.01]$, dashed double dotted lines represent the results for starch B with $[\overline{M}_n, \overline{M}_w, N, PD] = [4100, 5430, 22496, 1.324]$, while solid lines represent the results for starch C with $[\overline{M}_n, \overline{M}_w, N, PD] = [160, 212, 878, 1.325]$.

Finally, the SSF of starch with different distribution by the yeast which excretes both α -amylase and glucoamylase is shown in Figure 5.16. From the figure, despite the differences in the starch employed, surprisingly the resulting SSF profiles are generally indifferent. This implies that the presence of α -amylase is critical in the SSF of starch to neutralize the effect of different starch distributions and to generate sufficient non-reducing ends for glucoamylase to produce glucose. For pure enzymatic hydrolysis of starch using a mixture of α -amylase and glucoamylase, such a synergy between both enzymes had been observed in the literature (Fujii, Homma, & Taniguchi, 1988; Fujii & Kawamura, 1985). In fact, the results from Wong, Robertson, Lee, and Wagschal (2007) showed that the conversion profiles (in terms of glucose production) for the hydrolysis of corn starch and wheat starch by a combination of purified α -amylase and glucoamylase were dynamically similar, thus further corroborating the predictions. The generally indifferent SSF profiles further suggest that the yeast strain indeed has an optimal strategy for managing the excretion and the consumption machineries. Apparently the excretion and the consumption machineries of the yeast strain had been “ecologically tuned” to ensure optimum survival in responding to changes not only in the consumable substrates (DP1 – DP3), but also in the type of starch employed. Since α -amylase hydrolyzes starch randomly, the accumulation of the DP2 sugar in the broth is an indication of the generation of non-reducing ends (i.e. substrates for glucoamylase attack) due to hydrolysis by α -amylase. In this case, yeast cleverly utilizes its cellular resources by secreting a ‘sufficient’ amount of α -amylase at the first 50 h of fermentation such that the DP2 sugar (and hence the substrates for glucoamylase) is always in excess in the broth for all starches. Had lesser α -amylase been produced, there could have been insufficient substrates for glucoamylase attack (e.g. for starches A and B), resulting in the slower generation of glucose. On the other hand, within a finite pool of cellular resources, had more α -amylase been produced during this phase (and thus lesser of other enzymes, e.g.

glucoamylase, is produced), the reduced amount of glucoamylase would similarly limit the production of glucose and this would imply a wastage of cellular resources. It is indeed a pleasant surprise that the interlinked PBM-CM model is able to quantitatively demonstrate the synergistic benefits of dual enzyme secretion. With more experimental data, it is believed that the interlinked model could extract deeper insights from the complex SSF process.

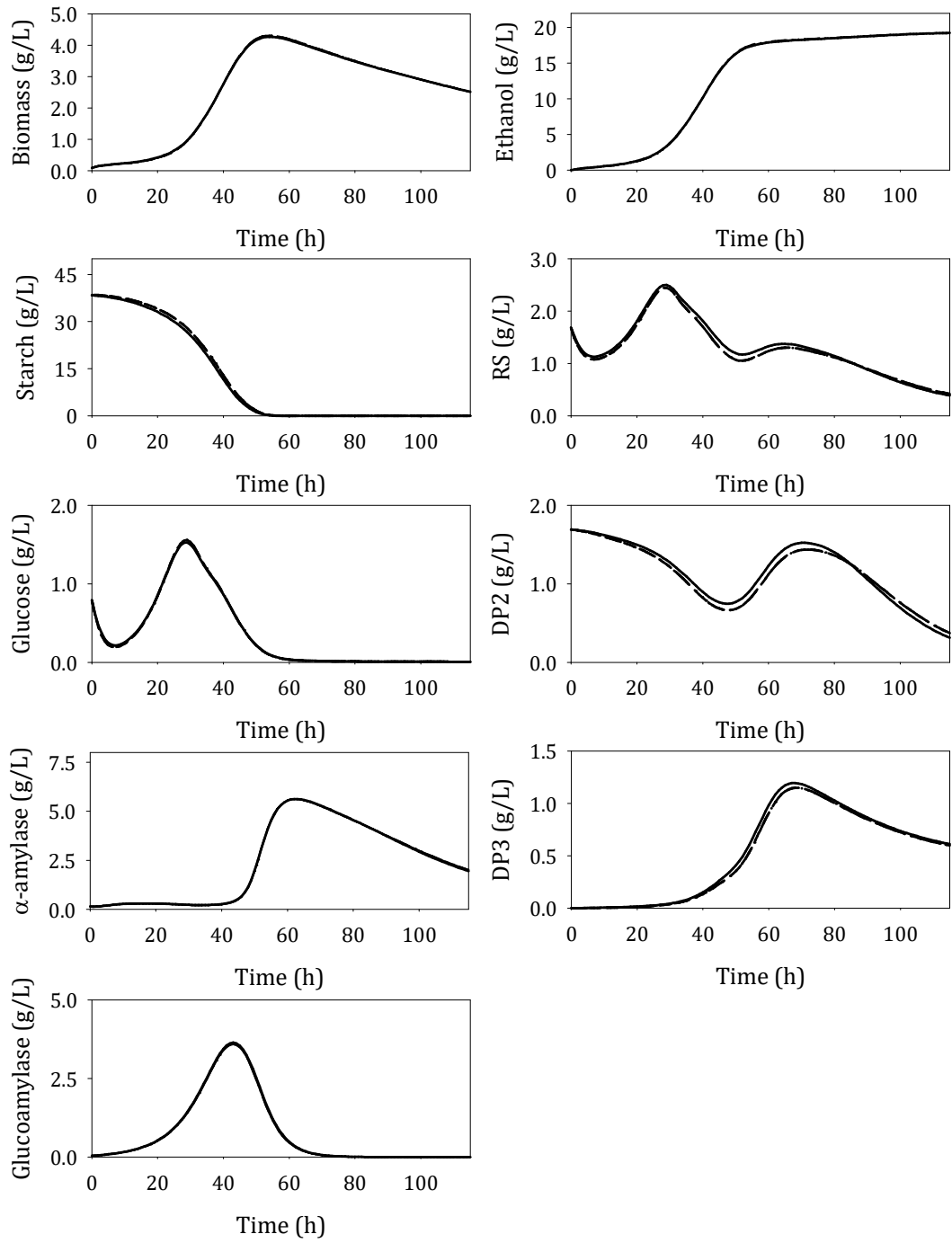


Figure 5.16: Transients of various quantities in the fermentation broth when the yeast is capable of excreting both α -amylase and glucoamylase. Here, dashed lines represent the results for starch A with $[\overline{M}_n, \overline{M}_w, N, PD] = [10000, 10100, 22496, 1.01]$, dashed double dotted lines represent the results for starch B with $[\overline{M}_n, \overline{M}_w, N, PD] = [4100, 5430, 22496, 1.324]$, while solid lines represent the results for starch C with $[\overline{M}_n, \overline{M}_w, N, PD] = [160, 212, 878, 1.325]$.

5.4 Concluding Remarks

The proposed methodology here is general and could be used to interlink the PBM and the CM to the preferred respective level of details, therefore presenting itself as a powerful tool for the systematic modelling of the interaction between the microbes and the environment containing complex substrates. It should be emphasized that the findings of the two case studies are dependent upon (i) the quality of the experimental data used for calibration, and (ii) the fidelity of abstracting both the enzymatic scission and the microbial assimilation of the resulting simple sugars. Robust experimental data on SSF with temporal evolution of the molecular distributions are a rare find, and certainly greater efforts are due. To facilitate model fine-tuning as raised in the second point, Table 5.6 summarizes the key assumptions that were made to construct the interlinked PBM and CM framework. If one desires further sophistication to the model (e.g. motivated by experimental discoveries), the simplifications outlined here can be altered accordingly without compromising the interlinking framework. Despite the relative simplicity of these assumptions, the case studies revealed that the proposed models exhibited features which could not be attained by the traditional approach of modelling SSF processes. As a result of the synergy between the PBM and the CM framework, the essential details of polymeric scission and cellular metabolism can be captured. Important properties of the polymer distribution, which are often critical to the design and operation of bioreactors with polymeric broths are also readily available.

Table 5.6: Key assumptions that can be relaxed without compromising the interlinked PBM and CM framework

No.	Assumption
1	All the carbon and energy sources are part of the polymeric substrates, i.e. $\mathcal{N} \subset \mathcal{S}$
2	Extracellular enzymes produced within the cells do not accumulate appreciably and are excreted instantaneously
3	Activities of the extracellular enzymes are not regulated
4	Starch is a linear population of polymers with DP greater than that of dextrin
5	Bounded sugars cannot be assimilated by yeast

CHAPTER 6 : CONCLUSIONS, THESIS AND RECOMMENDATION

6.1 Conclusions and Thesis

The ultimate aim of this work is to develop an interlinked PBM and CM framework for the SSF of natural polymers. To accomplish this, the necessary numerical techniques for solving PBEs were first established prior to laying out the framework for interlinking the PBM and the CM components. To systematically achieve the ultimate aim of this thesis, two research objectives were formulated and met, as summarized below:

- a) The first objective is to select or modify a potential numerical technique for approximating and solving PBEs for both chain-end and random scission as well as their combination thereof and also to explore the inherent characteristics of the resulting formulation.

Together with newly proposed guidelines to treat the lower molecular size range as a discrete domain in conjunction with a continuous domain in the upper ranges, the modified FP technique not only retains its original strengths in simulating both chain-end and random scission, but also captures accurately and efficiently the distribution of oligomers including monomers. From the results of the benchmarking process, the FP technique (implemented through the same discrete-continuous meshing strategy) was shown to be adept in predicting the simultaneously occurring random and chain-end scissions at only a fraction of the computational expense, even for problems with the size of N up to $\sim O(10^5)$. As had been done for chain-end scission, observations of the performance of the FP technique for random scission were also deliberated

and finally, the meshing guidelines for chain-end scission were revised accordingly for simulations which involve random scission.

- b) The second objective is to interlink the PBM and CM for modelling the batch growth of a microbial strain capable of simultaneously hydrolyzing a natural polymer and fermenting the resulting smaller saccharides. In addition, the resulting model should be used to analyze the growth of microbes on complex nutrients resulting from the individual or the combined actions of enzymes exhibiting random and chain-end scission behaviour.

This objective was achieved in that the general linkage between the PBM and the CM components was established in this study. A notable feature is the flexible linkage, which allows the individual PBM and CM models to be independently modified to the desired levels of detail. A more general treatment of the secretion of extracellular enzyme(s) was also proposed in the CM model. Through two case studies involving (i) chain-end and (ii) mixed chain-end and random scission enzymes, the calibrated interlinked model captured features not attainable by existing approaches. Other than being capable of capturing the effect of various enzymatic actions on the temporal evolution of the polymer distribution and how the microbes respond to the diverse oligomer environment due to the breakdown of starch, the interlinked PBM-CM model was able to do more. Specifically, in case study II the model was able to reveal that the excretion and the consumption machineries of the yeast strain had been “ecologically tuned” to ensure its optimum survival in responding to changes not only in the consumable substrates (DP1 – DP3) as advocated in the traditional CM framework but also the changes in the type of starch employed.

The evidence throughout this work supports the thesis that it is crucial to capture both the distribution of polymers as well as the regulatory nature of microbial metabolism when abstracting the dynamic complexities of the SSF of natural polymers. One possibility of achieving this is through the interlinking of the PBM and the CM components.

6.2 Recommendations and Future Work

Based on ideas developed in this work, several future extensions are possible:

- a) Instead of using the FP technique for simulating chain-end monomer scission, the formulation can be altered to capture the removal of an n -mer ($n = 2, 3, \dots$) from the end of a polymer chain. As it is important to maintain continuity between the discrete and the continuous region, the joint between the two regions for the case of $v_m \geq 2$ is currently unclear. This ability to simulate the removal of an n -mer from the end of a polymer chain is especially important as some exo-acting depolymerase (e.g. cellobiohydrolase I) had been reported to remove a dimer unit from the end of the polymer chain (Griggs et al., 2012a).
- b) The mixed scission models, suitably calibrated, can be used to study the optimum dosage required of the respective enzymes to achieve an intended blend of different sugars.
- c) Increased sophistication of the respective PBM and CM components can be used to expand the amount of information at the molecular level which can be gleaned from the framework.
- d) The interlinked PBM-CM framework should also be extended to deal with mixed microbial cultures. Such a mixed culture is not uncommon in SSF where one microbe

is responsible for excreting depolymerization enzymes while the other ferments the depolymerized components into desired products (Nakamura et al., 1997).

- e) A combination of some/all of the above can be done to further explore the potential of the framework. Moreover, instead of a pure starch system, the interlinked PBM-CM framework can be used to explore the effect of having multiple polymeric substrates (e.g. starch-cellulose mixture) on the dynamics of fermentation.
- f) Experimental investigations should be carried out to further validate the predictions from the interlinked model. One example is the effect of different starch distributions on the excretion and consumption behaviour of the microbes. This could only be probed by careful experimental studies.

REFERENCES

- Adam, A. C., Latorre - Garc í a, L., & Polaina, J. (2004). Structural analysis of glucoamylase encoded by the STA1 gene of *Saccharomyces cerevisiae* (var. diastaticus). *Yeast*, 21(5), 379-388.
- Aiken, R. C. (1985). *Stiff computation* (Vol. 169): Oxford University Press New York, Oxford.
- Alexopoulos, A. H., Pladis, P., & Kiparissides, C. (2013). Nonhomogeneous Mixing Population Balance Model for the Prediction of Particle Size Distribution in Large Scale Emulsion Polymerization Reactors. *Industrial & engineering chemistry research*, 52(35), 12285-12296. doi:10.1021/ie303500k
- Altintas, M. M., Kirdar, B., Onsan, Z. I., & Ulgen, K. O. (2002). Cybernetic modelling of growth and ethanol production in a recombinant *Saccharomyces cerevisiae* strain secreting a bifunctional fusion protein. *Process Biochemistry*, 37(12), 1439 - 1445.
- Anuradha, R., Suresh, A. K., & Venkatesh, K. V. (1999). Simultaneous saccharification and fermentation of starch to lactic acid. *Process Biochemistry*, 35(3), 367-375.
- Arga, K. Y., Çakır, T., Pir, P., Özer, N., Altıntaş, M. M., & Ülgen, K. (2004). Transfer function approach in structured modeling of recombinant yeast utilizing starch. *Process Biochemistry*, 39(10), 1237-1248.
- Atmuri, A. K., Henson, M. A., & Bhatia, S. R. (2013). A population balance equation model to predict regimes of controlled nanoparticle aggregation. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 436(0), 325-332.
- Attarakih, M. M., Bart, H. J., & Faqir, N. M. (2003). Optimal moving and fixed grids for the solution of discretized population balances in batch and continuous systems: Droplet breakage. *Chemical Engineering Science*, 58(7), 1251-1269.

- Azmi, A. S., Ngoh, G. C., Mel, M., & Hasan, M. (2010). Ragi tapai and *Saccharomyces cerevisiae* as potential coculture in viscous fermentation medium for ethanol production. *African Journal of Biotechnology*, 9(42), 7122 - 7127.
- Bailey, J. E., & Ollis, D. F. (1986). *Biochemical Engineering Fundamentals* (2nd ed. ed.). New York: McGraw-Hill Book Co.
- Bansal, P., Hall, M., Realff, M. J., Lee, J. H., & Bommarius, A. S. (2009). Modeling cellulase kinetics on lignocellulosic substrates. *Biotechnology advances*, 27(6), 833-848.
- Batistote, M., Cruz, S. H., & Ernandes, J. R. (2006). Altered patterns of maltose and glucose fermentation by brewing and wine yeasts influenced by the complexity of nitrogen source. *Journal of the Institute of Brewing*, 112(2), 84-91.
- Bernfeld, P. (1955). Amylases alpha and beta. *Methods Enzymol*, 1, 149 - 150.
- Besselink, T., Baks, T., Janssen, A. E. M., & Boom, R. M. (2008). A stochastic model for predicting dextrose equivalent and saccharide composition during hydrolysis of starch by α -amylase. *Biotechnology and Bioengineering*, 100(4), 684 - 697.
- Birol, G., Kirdar, B., & Önsan, Z. İ. (2002). A simple structured model for biomass and extracellular enzyme production with recombinant *Saccharomyces cerevisiae* YPB-G. *Journal of Industrial Microbiology & Biotechnology*, 29, 111 - 116.
- Birol, G., Önsan, Z. İ., Kirdar, B., & Oliver, S. G. (1998). Ethanol production and fermentation characteristics of recombinant *Saccharomyces cerevisiae* strains grown on starch. *Enzyme and Microbial Technology*, 22(8), 672-677.
- Brandam, C., Meyer, X. M., Proth, J., Strehaiano, P., & Pingaud, H. (2003). An original kinetic model for the enzymatic hydrolysis of starch during mashing. *Biochemical Engineering Journal*, 13, 43 - 52.
- Breuninger, W. F., Piyachomkwan, K., & Sriroth, K. (2009). Tapioca/Cassava Starch: Production and Use. In J. N. Bemiller & R. L. Whistler (Eds.), *Starch: Chemistry*

- and technology* (pp. 541 - 568). Burlington, MA: Food Science and Technology International Series, Academic Press
- Canu, P. (2005). Prediction of multimodal distributions in breakage processes. *Industrial & engineering chemistry research*, 44(8), 2649-2658.
- Caplice, E., & Fitzgerald, G. F. (1999). Food fermentations: role of microorganisms in food production and preservation. *International journal of food microbiology*, 50(1), 131-149.
- Chang, S.-Y., Delwiche, S. R., & Wang, N. S. (2002). Hydrolysis of wheat starch and its effect on the falling number procedure: mathematical model. *Biotechnology and Bioengineering*, 79(7), 768 - 775.
- Chavan, A. R., Raghunathan, A., & Venkatesh, K. V. (2009). Modeling and experimental studies on intermittend starch feeding and citrate addition in simultaneous saccharification and fermentation of starch into flavor compounds. *Journal of Industrial Microbiology and Biotechnology*, 36(4), 509 - 519.
- Cox, E. (2005). Chapter 9 - Fundamental Concepts of Genetic Algorithms. In E. Cox (Ed.), *Fuzzy Modeling and Genetic Algorithms for Data Mining and Exploration* (pp. 343-420). San Francisco: Morgan Kaufmann.
- Davis, R. A. (2008). Parameter estimation for simultaneous saccharification and fermentation of food waste into ethanol using Matlab Simulink *Biotechnology for Fuels and Chemicals* (pp. 379-389): Springer.
- De Moraes, L. M. P., Astolfi-Filho, S., & Oliver, S. G. (1995). Development of yeast strains for the efficient utilisation of starch: Evaluation of constructs that express α -amylase and glucoamylase separately or as bifunctional fusion proteins. *Applied microbiology and biotechnology*, 43(6), 1067-1076.

- Dean III, S. W., & Rollings, J. E. (1992). Analysis and quantification of a mixed exo-acting and endo-acting polysaccharide depolymerization system. *Biotechnology and Bioengineering*, 39, 968 - 976.
- Deng, Y., Li, S., Xu, Q., Gao, M., & Huang, H. (2012). Production of fumaric acid by simultaneous saccharification and fermentation of starchy materials with 2-deoxyglucose-resistant mutant strains of *Rhizopus oryzae*. *Bioresource Technology*, 107(0), 363-367.
- Detera, S. D., & Friedberg, F. (1979). Molecular weight of *B. subtilis* alpha-amylase derived from chemical studies. *Int J Pept Protein Res*, 14(4), 364-372.
- Diemer, R. B., & Olson, J. H. (2002). A moment methodology for coagulation and breakage problems: Part 1 - analytical solution of the steady-state population balance. *Chemical Engineering Science*, 57, 2193 - 2209.
- Dole, M. (1972). *The radiation chemistry of macromolecules* (Vol. I). New York: Academic Press.
- Duval, E. H., Alves Jr, S. L., Dunn, B., Sherlock, G., & Stambuk, B. U. (2010). Microarray karyotyping of maltose - fermenting *Saccharomyces* yeasts with differing maltotriose utilization profiles reveals copy number variation in genes involved in maltose and maltotriose utilization. *Journal of applied microbiology*, 109(1), 248-259.
- Eklund, R., & Zacchi, G. (1995). Simultaneous saccharification and fermentation of steam-pretreated willow. *Enzyme and Microbial Technology*, 17(3), 255-259.
- El-Zawawy, W. K., Ibrahim, M. M., Abdel-Fattah, Y. R., Soliman, N. A., & Mahmoud, M. M. (2011). Acid and enzyme hydrolysis to convert pretreated lignocellulosic materials into glucose for ethanol production. *Carbohydrate Polymers*, 84(3), 865-871.

- Ernandes, J. R., D'Amore, T., Russell, I., & Stewart, G. G. (1992). Regulation of glucose and maltose transport in strains of *Saccharomyces*. *Journal of industrial microbiology*, 9(2), 127-130.
- Eyre, D., Everson, R. C., & Campbell, Q. P. (1998). New parameterization for a discrete batch grinding equation. *Powder Technology*, 98(3), 265-272.
- Fagerström, R. (1991). Subsite mapping of *Hormoconis resinae* glucoamylases and their inhibition by gluconolactone. *Journal of General Microbiology*, 137, 1001 - 1008.
- Fierobe, H. P., Mirgorodskaya, E., Frandsen, T. P., Roepstorff, P., & Svensson, B. (1997). Overexpression and characterization of *Aspergillus awamori* wild-type and mutant glucoamylase secreted by the methylotrophic yeast *Pichia pastoris*: comparison with wild-type recombinant glucoamylase produced using *Saccharomyces cerevisiae* and *Aspergillus niger* as hosts. *Protein Expr Purif*, 9(2), 159-170.
- Fujii, M., Homma, T., & Taniguchi, M. (1988). Synergism of α -amylase and glucoamylase on hydrolysis of native starch granules. *Biotechnology and Bioengineering*, 32, 910 - 915.
- Fujii, M., & Kawamura, Y. (1985). Synergistic action of α -amylase and glucoamylase on hydrolysis of starch. *Biotechnology and Bioengineering*, XXVII, 260 - 265.
- Gadgil, C. J., Bhat, P. J., & Venkatesh, K. V. (1996). Cybernetic model for the growth of *Saccharomyces cerevisiae* on melibiose. *Biotechnology Progress*, 12(6), 744-750.
- Gelbard, F., & Seinfeld, J. H. (1979). The general dynamic equation for aerosols: Theory and application to aerosol formation and growth. *Journal of Colloid and Interface Science*, 68(2), 363 - 382.
- Geng, J., Song, H.-S., Yuan, J., & Ramkrishna, D. (2012). On enhancing productivity of bioethanol with multiple species. *Biotechnology and Bioengineering*, 109(6).

- Glick, B. R. (1995). Metabolic load and heterologous gene-expression. *Biotechnology advances*, 13(2), 247-261.
- Griggs, A. J., Stickel, J. J., & Lischeske, J. J. (2012a). A mechanistic model for enzymatic saccharification of cellulose using continuous distribution kinetics I: Depolymerization by EG_I and CBH_I. *Biotechnology and Bioengineering*, 109(3), 665 - 675.
- Griggs, A. J., Stickel, J. J., & Lischeske, J. J. (2012b). A mechanistic model for enzymatic saccharification of cellulose using continuous distribution kinetics II: Cooperative enzyme action, solution kinetics, and product inhibition. *Biotechnology and Bioengineering*, 109(3), 676 - 685.
- Hamilton, R. A., Curtis, J. S., & Ramkrishna, D. (2003). Beyond log - normal distributions: Hermite spectra for solving population balances. *AIChE Journal*, 49(9), 2328-2343.
- Hetényi, K., Németh, Á., & Sevela, B. (2011). Investigation and modeling of lactic acid fermentation on wheat starch via SSF, CHF and SHF technology. *Chemical Engineering*, 55(1), 11-16.
- Hiemenz, P. C., & Lodge, T. P. (2007). *Polymer Chemistry* (2nd ed.). Boca Raton, Florida: CRC Press.
- Hill, P. J., & Ng, K. M. (1995). New discretization procedure for the breakage equation. *AIChE Journal*, 41, 1204 - 1216.
- Hiromi, K. (1970). Interpretation of dependency of rate parameters on the degree of polymerization of substrate in enzyme-catalyzed reactions. Evaluation of subsite affinities of exo-enzyme. *Biochemical and Biophysical Research Communications*, 40(1), 1 - 6.

- Hiromi, K., Nitta, Y., Numata, C., & Ono, S. (1973). Subsite affinities of glucoamylase: Examination of the validity of the subsite theory. *Biochimica et Biophysica Acta*, 302, 362 - 375.
- Hofvendahl, K., Åkerberg, C., Zacchi, G., & Hahn-Hägerdal, B. (1999). Simultaneous enzymatic wheat starch saccharification and fermentation to lactic acid by *Lactococcus lactis*. *Applied microbiology and biotechnology*, 52(2), 163-169.
- Hosseini, S. A., & Shah, N. (2011a). Enzymatic hydrolysis of cellulose part II: Population balance modelling of hydrolysis by exoglucanase and universal kinetic model. *Biomass and Bioenergy*, 35, 3830 - 3840.
- Hosseini, S. A., & Shah, N. (2011b). Modelling enzymatic hydrolysis of cellulose part I: Population balance modelling of hydrolysis by endoglucanase. *Biomass and Bioenergy*, 35, 3841 - 3848.
- Houcque, D. (2008). Applications of MATLAB: Ordinary differential equations (ODE). *Northwestern University*.
- Iwasa, S., Aoshima, H., Hiromi, K., & Hatano, H. (1974). Subsite affinities of bacterial liquefying α -amylase evaluated from the rate parameters of linear substrates. *Journal of Biochemistry*, 75, 969 - 978.
- Jang, M.-F., & Chou, S.-H. (2013). Modeling and optimization of bioethanol production via a simultaneous saccharification and fermentation process using starch. *Journal of Chemical Technology and Biotechnology*, 88(6), 1164 - 1174.
- John, R. P., Nampoothiri, K. M., & Pandey, A. (2009). Direct lactic acid fermentation: Focus on simultaneous saccharification and lactic acid production. *Biotechnology advances*, 27(2), 145-152.
- John, V., Angelov, I., Oncul, A. A., & Thevenin, D. (2007). Techniques for the reconstruction of a distribution from a finite number of its moments. *Chemical Engineering Science*, 62, 2890 - 2904.

- Karimi, K., Emtiazi, G., & Taherzadeh, M. J. (2006). Ethanol production from dilute-acid pretreated rice straw by simultaneous saccharification and fermentation with *Mucor indicus*, *Rhizopus oryzae*, and *Saccharomyces cerevisiae*. *Enzyme and Microbial Technology*, 40(1), 138-144.
- Karmore, V., & Madras, G. (2001). Degradation kinetics for polymer mixtures in solution. *Industrial and Engineering Chemistry Research*, 40, 1306 - 1311.
- Kearsley, M. W., & Dziedzic, S. Z. (1995). *Handbook of starch hydrolysis products and their derivatives*. London: Blackie Academic.
- Ko, J., Su, W.-J., Chien, I.-L., Chang, D.-M., Chou, S.-H., & Zhan, R.-Y. (2010). Dynamic modeling and analyses of simultaneous saccharification and fermentation process to produce bio-ethanol from rice straw. *Bioprocess and Biosystems Engineering*, 33(2), 195 - 205.
- Kobayashi, F., & Nakamura, Y. (2003). Mathematical model for growth process of a recombinant yeast having saccharification and fermentation activities. *Journal of Chemical Technology and Biotechnology*, 78(9), 985-994.
- Kobayashi, F., & Nakamura, Y. (2004). Mathematical model of direct ethanol production from starch in immobilized recombinant yeast culture. *Biochemical Engineering Journal*, 21(1), 93-101.
- Koljonen, T., Hämäläinen, J. J., Sjöholm, K., & Pietilä, K. (1995). A model for the prediction of fermentable sugar concentrations during mashing. *Journal of Food Engineering*, 26, 329 - 350.
- Kompala, D. S., & Ramkrishna, D. (1986). Investigation of bacterial growth on mixed substrates : Experimental evaluation of cybernetic models. *Biotechnology and Bioengineering*, 28, 1044 - 1055.

- Kompala, D. S., Ramkrishna, D., & Tsao, G. T. (1984). Cybernetic modeling of microbial growth on multiple substrates. *Biotechnology and Bioengineering*, 26(11), 1272-1281.
- Kondo, H., Nakatani, H., Hiromi, K., & Matsuno, R. (1978). Multiple attack in porcine pancreatic α -amylase-catalyzed hydrolysis of amylose studied with a fluorescence probe. *Journal of Biochemistry*, 84(2), 403 - 417.
- Kondo, H., Nakatani, H., Matsuno, R., & Hiromi, K. (1980). Product distribution in amylase-catalyzed hydrolysis of amylose: comparison of experimental results with theoretical predictions. *Journal of Biochemistry*, 87(4), 1053 - 1070.
- Kostoglou. (2000). Mathematical analysis of polymer degradation with chain-end scission. *Chemical Engineering Science*, 55, 2507 - 2513.
- Kostoglou, M. (2007). Extended cell average technique for the solution of coagulation equation. *Journal of Colloid and Interface Science*, 306(1), 72 - 81.
- Kostoglou, M., & Karabelas, A. J. (2002). An assessment of low-order methods for solving the breakage equation. *Powder Technology*, 127, 116 - 127.
- Kostoglou, M., & Karabelas, A. J. (2004). Optimal low order methods of moments for solving the fragmentation equation. *Powder Technology*, 143 - 144, 280 - 290.
- Kostoglou, M., & Karabelas, A. J. (2009). On sectional techniques for the solution of the breakage equation. *Computers and Chemical Engineering*, 33, 112 - 121.
- Koutinas, A. A., Wang, R., Kookos, I. K., & Webb, C. (2003). Kinetic parameters of *Aspergillus awamori* in submerged cultivations on whole wheat flour under oxygen limiting conditions. *Biochemical Engineering Journal*, 16, 23 - 34.
- Kroumov, A. D., Módenes, A. N., & de Araujo Tait, M. C. (2006). Development of new unstructured model for simultaneous saccharification and fermentation of starch to ethanol by recombinant strain. *Biochemical Engineering Journal*, 28, 243 - 255.

- Kumar, J., Peglow, M., Warnecke, G., & Heinrich, S. (2008). An efficient numerical technique for solving population balance equation involving aggregation, breakage, growth and nucleation. *Powder Technology*, 182, 81 - 104.
- Kumar, S., & Ramkrishna, D. (1996a). On the solution of population balance equations by discretization - I. A fixed pivot technique. *Chemical Engineering Science*, 51(8), 1311 - 1332.
- Kumar, S., & Ramkrishna, D. (1996b). On the solution of population balance equations by discretization - II. A moving pivot technique. *Chemical Engineering Science*, 51(8), 1333 - 1342.
- Kurasin, M., & Valjamae, P. (2011). Processivity of celobiohydrolase is limited by the substrate. *Journal of Biological Chemistry*, 286(1), 169 - 177.
- Kusunoki, K., Kawakami, K., Shiraishi, F., Kato, K., & Kai, M. (1982). A kinetic expression for hydrolysis of soluble starch by glucoamylase. *Biotechnology and Bioengineering*, XXIV, 347 - 354.
- Lee, C. G., Kim, C. H., & Rhee, S. K. (1992). A kinetic model and simulation of starch saccharification and simultaneous ethanol fermentation by amyloglucosidase and *Zymomonas mobilis*. *Bioprocess Engineering*, 7, 335 - 341.
- Lee, Y.-S., Lee, W. G., Chang, Y. K., & Chang, H. N. (1995). Modelling of ethanol production by *Saccharomyces cerevisiae* from a glucose and maltose mixture. *Biotechnology Letters*, 17(8), 791 - 796.
- Liu, Y., & Tadé, M. O. (2004). New wavelet-based adaptive method for the breakage equation. *Powder Technology*, 139(1), 61-68.
- Madras, G., Chung, G. Y., Smith, J. M., & McCoy, B. J. (1997). Molecular weight effect on the dynamics of polystyrene degradation. *Industrial and Engineering Chemistry Research*, 36, 2019 - 2024.

- Madras, G., Smith, J. M., & McCoy, B. J. (1996a). Degradation of poly(methyl methacrylate) in solution. *Industrial and Engineering Chemistry Research*, 35, 1795 - 1800.
- Madras, G., Smith, J. M., & McCoy, B. J. (1996b). Thermal degradation of poly(α -methylstyrene) in solution. *Polymer Degradation and Stability*, 52, 349 - 358.
- Madras, G., Smith, J. M., & McCoy, B. J. (1997). Thermal degradation kinetics of polystyrene in solution. *Polymer Degradation and Stability*, 58, 131 - 138.
- Mantzaris, N. V. (2005). Transient and asymptotic behaviour of the binary breakage problem. *Journal of Physics A: Mathematical and General*, 38(23), 5111.
- Marc, A., & Engasser, J. M. (1983). A kinetic model of starch hydrolysis by α - and β -amylase during mashing. *Biotechnology and Bioengineering*, XXV, 481 - 496.
- Marchal, L. M., Ulijin, R. V., de Gooijer, C. D., Franke, G. T., & Tramper, J. (2003). Monte Carlo simulation of the α -amylolysis of amylopectin potato starch. 2. α -amylolysis of amylopectin. *Bioprocess and Biosystem Engineering*, 26, 123 - 132.
- Marchal, L. M., Zondervan, J., Bergsma, J., Beftink, H. H., & Tramper, J. (2001). Monte Carlo simulation of the α -amylolysis of amylopectin potato starch. 1. modeling of the structure of amylopectin. *Bioprocess and Biosystem Engineering*, 24, 163 - 170.
- Matsumura, M., & Hirata, J. (1989). Continuous simultaneous saccharification and fermentation of raw starch in a membrane reactor. *Journal of Chemical Technology and Biotechnology*, 46(4), 313-326.
- McCoy, B. J. (1999). Distribution kinetics for temperature-programmed pyrolysis. *Industrial and Engineering Chemistry Research*, 38, 4531 - 4537.
- McCoy, B. J. (2001). Polymer thermogravimetric analysis: effects of chain-end and reversible random scission. *Chemical Engineering Science*, 56, 1525 - 1529.

- McCoy, B. J., & Madras, G. (2001). Discrete and continuous models for polymerization and depolymerization. *Chemical Engineering Science*, 56, 2831 - 2836.
- McCoy, B. J., & Wang, M. (1994). Continuous-mixture fragmentation kinetics: Particle size reduction and molecular cracking. *Chemical Engineering Science*, 49(22), 3773 - 3785.
- McMillan, J. D., Newman, M. M., Templeton, D. W., & Mohagheghi, A. (1999). *Simultaneous saccharification and cofermentation of dilute-acid pretreated yellow poplar hardwood to ethanol using xylose-fermenting Zymomonas mobilis*. Paper presented at the Twentieth Symposium on Biotechnology for Fuels and Chemicals.
- Miranda, M., & Murado, M. A. (1991). Mass transfer control of enzymatic hydrolysis of polysaccharides by glucoamylase. *Enzyme and Microbial Technology*, 13, 142 - 147.
- Monod, J. (1942). *Recherches sur la croissance des cultures bacteriennes*. Paris, France: Hermann et Cie.
- Montesinos, T., & Navarro, J.-M. (2000). Production of alcohol from raw wheat flour by amyloglucosidase and *Saccharomyces cerevisiae*. *Enzyme and Microbial Technology*, 27, 362 - 370.
- Montesinos, T., & Navarro, J.-M. (2000). Role of the maltose in the simultaneous-saccharification-fermentation process from raw wheat starch and *Saccharomyces cerevisiae*. *Bioprocess Engineering*, 23(4), 319-322.
- Morales-Rodriguez, R., Gernaey, K. V., Meyer, A. S., & Sin, G. (2011). A Mathematical Model for Simultaneous Saccharification and Co-fermentation (SSCF) of C6 and C5 Sugars. *Chinese Journal of Chemical Engineering*, 19(2), 185-191.
- Morales-Rodriguez, R., Meyer, A. S., Gernaey, K. V., & Sin, G. (2011). Dynamic model-based evaluation of process configurations for integrated operation of hydrolysis

- and co-fermentation for bioethanol production from lignocellulose. *Bioresource Technology*, 102(2), 1174-1184.
- Murthy, G. S., Johnston, D. B., Rausch, K. D., Tumbleson, M. E., & Singh, V. (2012). A simultaneous saccharification and fermentation model for dynamic growth environments. *Bioprocess and Biosystems Engineering*, 35(4), 519-534.
- Murthy, G. S., Johnston, D. B., Rauseh, K. D., Tumbleson, M. E., & Singh, V. (2011). Starch hydrolysis modeling: application to fuel ethanol production. *Bioprocess and Biosystem Engineering*, 34, 879 - 890.
- Nakamura, Y., Kobayashi, F., Ohnaga, M., & Sawada, T. (1997). Alcohol fermentation of starch by a genetic recombinant yeast having glucoamylase activity. *Biotechnology and Bioengineering*, 53(1), 21-25.
- Nakasaki, K., Murai, T., & Akiyama, T. (1988). Kinetic modeling of simultaneous saccharification and fermentation of cellulose. *Journal of chemical engineering of Japan*, 21(4), 436-438.
- Nopens, I., Beheydt, D., & Vanrolleghem, P. A. (2005). Comparison and pitfalls of different discretised solution methods for population balance models: a simulation study. *Computers and Chemical Engineering*, 29, 367 - 377.
- Ochoa, S., Yoo, A., Repke, J.-U., Wozny, G., & Yang, D. R. (2007). Modeling and parameter identification of the simultaneous saccharification-fermentation process for ethanol production. *Biotechnology Progress*, 23, 1454 - 1462.
- Oyerokun, F. T., & Vaia, R. A. (2012). Distribution in the grafting density of end-functionalized polymer chains adsorbed onto nanoparticle surfaces. *Macromolecules*, 45, 7649 - 7659.
- Paolucci-Jeanjean, D., Belleville, M., Rios, G. M., & Zakhia, N. (2000). Kinetics of continuous starch hydrolysis in a membrane reactor. *Biochemical Engineering Journal*, 6, 233 - 238.

- Paolucci-Jeanjean, D., Belleville, M., Zakhia, N., & Rios, G. M. (2000). Kinetics of cassava starch hydrolysis with Termamyl[®] Enzyme. *Biotechnology and Bioengineering*, 68(1), 71 - 77.
- Philippidis, G. P., Spindler, D. D., & Wyman, C. E. (1992). Mathematical modeling of cellulose conversion to ethanol by the simultaneous saccharification and fermentation process. *Applied Biochemistry and Biotechnology*, 34(1), 543-556.
- Podkaminer, K. K., Shao, X., Hogsett, D. A., & Lynd, L. R. (2011). Enzyme inactivation by ethanol and development of a kinetic model for thermophilic simultaneous saccharification and fermentation at 50 °C with *Thermoanaerobacterium saccharolyticum* ALK2. *Biotechnology and Bioengineering*, 108(6), 1268-1278.
- Polakovič, M., & Bryjak, J. (2004). Modelling of potato starch saccharification by an *Aspergillus niger* glucoamylase. *Biochemical Engineering Journal*, 18, 57 - 63.
- Presečki, A. V., Blažević, Z. F., & Vasić-Rački, D. (2013). Complete starch hydrolysis by the synergistic action of amylase and glucoamylase: impact of calcium ions. *Bioprocess Biosyst Eng*, 36, 1555 - 1562.
- Pretorius, I. S. (2000). Tailoring wine yeast for the new millennium: novel approaches to the ancient art of winemaking. *Yeast*, 16(8), 675-729.
- Ramkrishna, D. (1983). A cybernetic perspective of microbial growth *Foundations of Biochemical Engineering* (pp. 161 - 178). Washington, DC: American Chemical Society.
- Ramkrishna, D. (2000). *Population balances: Theory and applications to particulate systems in engineering*. San Diego, California, USA: Academic Press.
- Ramkrishna, D., & Singh, M. R. (2014). Population Balance Modeling: Current Status and Future Prospects. *Annual Review of Chemical and Biomolecular Engineering*, 5(1), 123-146.

- Ramkrishna, D., & Song, H.-S. (2012). Dynamic models of metabolism: Review of the cybernetic approach. *AIChE Journal*, 58(4), 986 - 997.
- Rangarajan, P., Bhattacharyya, D., & Grulke, E. (1998). HDPE liquefaction: Random chain scission model. *Journal of Applied Polymer Science*, 70, 1239 - 1251.
- Robyt, J. F. (2009). Enzymes and their action on starch. In J. N. Bemiller & R. L. Whistler (Eds.), *Starch: Chemistry and technology* (pp. 237 - 292). Burlington, MA: Food Science and Technology International Series, Academic Press.
- Robyt, J. F., & French, D. (1967). Multiple attack hypothesis of α -amylase action: Action of porcine pancreatic, human salivary, and *Aspergillus oryzae* α -amylases. *Archives of Biochemistry and Biophysics*, 122(1), 8 - 16.
- Rohatgi, A. (2012). WebPlotDigitizer: HTML5 based online tool to extract numerical data from plot images (Version 3.6). Retrieved from <http://arohatgi.info/WebPlotDigitizer/app/>
- Rothstein, D. M., Devlin, P. E., & Cate, R. L. (1986). Expression of alpha-amylase in *Bacillus licheniformis*. *J Bacteriol*, 168(2), 839-842.
- Saqib, A. A. N., & Whitney, P. J. (2011). Differential behaviour of the dinitrosalicylic acid (DNS) reagent towards mono- and di-saccharide sugars. *Biomass and Bioenergy*, 35(11), 4748-4750.
- Shen, J., & Agblevor, F. A. (2010). Modeling semi-simultaneous saccharification and fermentation of ethanol production from cellulose. *Biomass and Bioenergy*, 34(8), 1098 - 1107.
- Shen, J., & Agblevor, F. A. (2011). Ethanol production of semi-simultaneous saccharification and fermentation from mixture of cotton gin waste and recycled paper sludge. *Bioprocess and Biosystems Engineering*, 34(1), 33-43.
- Shin, D., Yoo, A., Kim, S. W., & Yang, D. R. (2006). Cybernetic modeling of simultaneous saccharification and fermentation for ethanol production from

- steam-exploded wood with *Brettanomyces custersii*. *Journal of Microbiology and Biotechnology*, 16(9), 1355-1361.
- Shiraishi, F., Kawakami, K., & Kusunoki, K. (1985). Kinetics of condensation of glucose into maltose and isomaltose in hydrolysis of starch by glucoamylase. *Biotechnology and Bioengineering*, XXVII, 498 - 502.
- Shuler, M. L., & Kargi, F. (2002). *Bioprocess Engineering: Basic Concepts*. Upper Saddle River, New Jersey: Prentice Hall.
- Sierks, M. R. (1988). *Mutagenesis of the active site of glucoamylase from Aspergillus awamori*. (Ph.D.), Iowa State University.
- Song, H. S., Morgan, J. A., & Ramkrishna, D. (2009). Systematic development of hybrid cybernetic models: Application to recombinant yeast co - consuming glucose and xylose. *Biotechnology and Bioengineering*, 103(5), 984-1002.
- Song, H. S., & Ramkrishna, D. (2011). Cybernetic models based on lumped elementary modes accurately predict strain - specific metabolic function. *Biotechnology and Bioengineering*, 108(1), 127-140.
- Staggs, J. E. J. (2002). Modelling random scission of linear polymers. *Polymer Degradation and Stability*, 76, 37 - 44.
- Staggs, J. E. J. (2004). Modelling end-chain scission and recombination of linear polymers. *Polymer Degradation and Stability*, 85, 759 - 767.
- Staggs, J. E. J. (2005). A continuous model for vapourisation of linear polymers by random scission and recombination. *Fire Safety Journal*, 40, 610 - 627.
- Staggs, J. E. J. (2006). Discrete bond-weighted random scission of linear polymers. *Polymer*, 47, 897 - 906.
- Sterling, W. J., & McCoy, B. J. (2001). Distribution kinetics of thermolytic macromolecular reactions. *AIChE Journal*, 47, 2289 - 2303.

- Steverson, E. M., Korus, R. A., Admassu, W., & Heimsch, R. C. (1984). Kinetics of the amylase system of *Saccharomycopsis fibulinger*. *Enzyme and Microbial Technology*, 6, 549 - 554.
- Stickel, J. J., & Griggs, A. J. (2012). Mathematical modeling of chain-end scission using continuous distribution kinetics. *Chemical Engineering Science*, 68, 656 - 659.
- Striegel, A. M. (2003). Determining and correcting "moment bias" in gradient polymer elution chromatography. *Journal of Chromatography A*, 996, 45 - 51.
- Sun, Y., & Cheng, J. (2002). Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology*, 83(1), 1-11.
- Taherzadeh, M. J., & Karimi, K. (2007). Enzymatic-based hydrolysis processes for ethanol from lignocellulosic materials: A review. *BioResources*, 2(4), 707-738.
- Thomson, J. M., Gaucher, E. A., Burgan, M. F., De Kee, D. W., Li, T., Aris, J. P., & Benner, S. A. (2005). Resurrecting ancestral alcohol dehydrogenases from yeast. *Nature Genetics*, 37(6), 630-635.
- Turner, B. G., Ramkrishna, D., & Jansen, N. B. (1989). Cybernetic modeling of bacteriol cultures at low growth rates: Single - substrate systems. *Biotechnology and Bioengineering*, 34(2), 252-261.
- Ülgen, K. Ö., Saygılı, B., Önsan, Z. İ., & Kırđar, B. (2002). Bioconversion of starch into ethanol by a recombinant *Saccharomyces cerevisiae* strain YPG-AB. *Process Biochemistry*, 37(10), 1157-1168.
- van Maris, A. J. A., Abbott, D. A., Bellissimi, E., van den Brink, J., Kuyper, M., Luttik, M. A. H., . . . Pronk, J. T. (2006). Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Antonie Van Leeuwenhoek*, 90(4), 391-418.

- van Zyl, J. M., van Rensburg, E., van Zyl, W. H., Harms, T. M., & Lynd, L. R. (2011). A kinetic model for simultaneous saccharification and fermentation of Avicel with *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, 108(4), 924-933.
- Vanni, M. (2000). Approximate population balance equations for aggregation-breakage processes. *Journal of Colloid and Interface Science*, 221, 143-160.
- Varner, J., & Ramkrishna, D. (1998). Application of cybernetic models to metabolic engineering: Investigation of storage pathways. *Biotechnology and Bioengineering*, 58(2 - 3), 282 - 291.
- Varner, J., & Ramkrishna, D. (1999a). Metabolic engineering from a cybernetic perspective. 1. Theoretical Preliminaries. *Biotechnology Progress*, 15(3), 407-425.
- Varner, J., & Ramkrishna, D. (1999b). The non-linear analysis of cybernetic models. Guidelines for model formulation. *Journal of Biotechnology*, 71(1), 67-103.
- Wang, M., Smith, J. M., & McCoy, B. J. (1995). Continuous kinetics for thermal degradation of polymer in solution. *AIChE Journal*, 41(6), 1521 - 1533.
- Wojciechowski, P. M., Koziol, A., & Noworyta, A. (2001). Iteration model of starch hydrolysis by amylolytic enzymes. *Biotechnology and Bioengineering*, 75(5), 530 - 539.
- Wong, D. W. S., Robertson, G. H., Lee, C. C., & Wagschal, K. (2007). Synergistic action of recombinant α -amylase and glucoamylase on the hydrolysis of starch granules. *The Protein Journal*, 26(3), 159 - 164.
- Young, J. D., & Ramkrishna, D. (2007). On the matching and proportional laws of cybernetic models. *Biotechnology Progress*, 23(1), 83-99.
- Zhang, J., Shao, X., Townsend, O. V., & Lynd, L. R. (2009). Simultaneous saccharification and co - fermentation of paper sludge to ethanol by *Saccharomyces cerevisiae* RWB222—Part I: Kinetic modeling and parameters. *Biotechnology and Bioengineering*, 104(5), 920-931.

- Ziff, R. M. (1992). An explicit solution to a discrete fragmentation model. *Journal of Physics A: Mathematical and General*, 25, 2569 - 2576.
- Ziff, R. M., & McGrady, E. D. (1985). The kinetics of cluster fragmentation and depolymerisation. *Journal of Physics A: Mathematical and General*, 18, 3027 - 3037.

LIST OF PUBLICATIONS AND CONFERENCES ATTENDED

Publications:

- Ho, Y. K., Doshi, P., Yeoh, H. K., & Ngoh, G. C. (2014). Modelling chain-end scission using the fixed pivot technique. *Chemical Engineering Science*, 116, 601, 601-610. DOI: 10.1016/j.ces.2014.05.035
- Ho, Y. K., Doshi, P., Yeoh, H. K., & Ngoh, G. C. (2015). Interlinked population balance and cybernetic models for the simultaneous saccharification and fermentation of natural polymers. *Biotechnology and Bioengineering*, 112, 112, 2084-2105. DOI: 10.1002/bit.25616
- Ho, Y. K., Doshi, P., Yeoh, H. K., & Ngoh, G. C. (2015). Why are two enzymes better than one for the efficient simultaneous saccharification and fermentation (SSF) of natural polymers? Hints from inside and outside a yeast. *Industrial & Engineering Chemistry Research*. DOI: 10.1021/acs.iecr.5b01667
- Ho, Y. K., Yeoh, H. K., & Mjalli, F. S. (2014). Generalized predictive control algorithm with real-time simultaneous modeling and tuning. *Industrial & Engineering Chemistry Research*, 53, 22, 9411-9426. DOI: 10.1021/ie401905w.

Conference Attended:

- Ho, Y. K., Yeoh, H. K., Ngoh, G. C., Mohamad Nor, M. I. (2014). Modelling enzymatic hydrolysis of starch by α -amylase. Accepted for oral presentation in the 6th Regional Conference on Chemical Engineering (RCChE) on 2 – 3 December 2013 in Manila, Philippines.

APPENDIX A: MODELLING CHAIN-END SCISSION

A.1 Theoretical Preliminaries of The Fixed Pivot Technique (S. Kumar and Ramkrishna, 1996a)

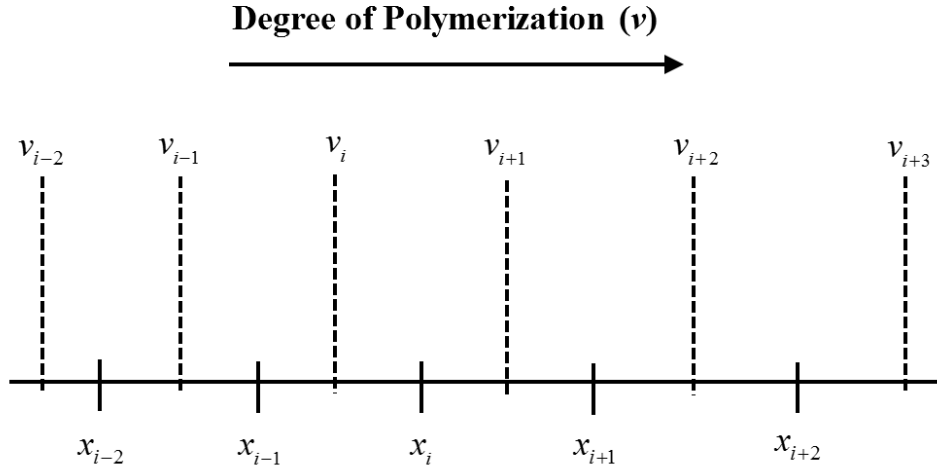


Figure A.1: Discretizing the DP axis into different intervals enclosed by $[v_i, v_{i+1}]$ where x_i is the representative pivot for the i -th interval and $v_i = (x_i + x_{i-1})/2$.

Consider the following continuous PBE for breakage:

$$\frac{\partial c(v, t)}{\partial t} = \int_v^\infty k(w) b(v, w) c(w, t) dw - k(v) c(v, t) \quad (\text{A.1})$$

The sectional method for solving PBE begins by discretizing the v -axis into intervals enclosed by $[v_i, v_{i+1}]$, as illustrated in Figure A.1. In the FP technique, a fixed representative size is chosen for each interval and is referred to as the pivot point (x_i). Within a particular interval, it is assumed that the molar concentration density is

concentrated at the pivot point and is zero elsewhere, i.e. $c(v, t) = \sum_{j=1}^{\infty} C_j(t) \delta(v - x_j)$. If

a polymer of size v falls between $[x_i, x_{i+1}]$, because the polymer does not fall exactly at the pivots x_i or x_{i+1} , the FP technique assumes that such a polymer be assigned to x_i and

x_{i+1} such that two integral properties of interest for the polymer populations be preserved.

This is done by assigning fractions $K(v, x_i)$ and $L(v, x_{i+1})$ to the polymer populations at x_i and x_{i+1} . For the preservation of two polymer properties $f_1(v)$ and $f_2(v)$, the fractional assignment can be written as:

$$K(v, x_i) f_1(x_i) + L(v, x_{i+1}) f_1(x_{i+1}) = f_1(v) \quad (\text{A.2})$$

$$K(v, x_i) f_2(x_i) + L(v, x_{i+1}) f_2(x_{i+1}) = f_2(v) \quad (\text{A.3})$$

Solving for fractions $K(v, x_i)$ and $L(v, x_{i+1})$:

$$K(v, x_i) = \frac{f_1(x_{i+1}) f_2(v) - f_1(v) f_2(x_{i+1})}{f_1(x_{i+1}) f_2(x_i) - f_1(x_i) f_2(x_{i+1})} \quad (\text{A.4})$$

$$L(v, x_{i+1}) = \frac{f_1(x_i) f_2(v) - f_1(v) f_2(x_i)}{f_1(x_i) f_2(x_{i+1}) - f_1(x_{i+1}) f_2(x_i)} \quad (\text{A.5})$$

The molar concentration of the i -th pivot can then be obtained by integrating the molar concentration density over the interval $[v_i, v_{i+1}]$:

$$\frac{dC_i(t)}{dt} = \frac{\partial}{\partial t} \int_{v_i}^{v_{i+1}} c(v, t) dv = \int_{v_i}^{v_{i+1}} \int_v^{\infty} k(w) b(v, w) c(w, t) dw dv - \int_{v_i}^{v_{i+1}} k(v) c(v, t) dv \quad (\text{A.6})$$

The FP procedure rewrites Eq. (A.6) in the following form:

$$\frac{dC_i(t)}{dt} = B - D \quad (\text{A.7})$$

where the death term, D is given by:

$$D = \int_{v_i}^{v_{i+1}} k(v) \sum_{j=1}^{\infty} C_j(t) \delta(v - x_j) dv = k_i C_i(t) \quad (\text{A.8})$$

As the death term for dC_i/dt does not interact with other intervals, there is no need for fractional assignments elaborated above.

For the birth term, because of the inter-interval interaction, i.e. by receiving from the scission of larger polymers, here the fractional assignment as discussed above comes into play. As only polymers that fall within $[v_i, v_{i+1}]$ matters, these are covered by the pivot intervals of $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$. Therefore, applying the principles shown in Eqs. (A.2) – (A.5), the following fractional assignments arise: $K(v, x_{i-1})$, $L(v, x_i)$, $K(v, x_i)$, and $L(v, x_{i+1})$. Since the current task is to write the birth term (B) for x_i , polymers that fall within the intervals of $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$ are respectively assigned to x_i through $L(v, x_i)$ and $K(v, x_i)$. Therefore, the birth term B can then be obtained in the following manner:

$$\begin{aligned}
B &= \int_{v_i}^{v_{i+1}} \int_v^\infty k(w) b(v, w) \sum_{j=1}^\infty C_j(t) \delta(w - x_j) dw dv \\
&= \int_{x_i}^{x_{i+1}} K(v, x_i) \int_v^\infty k(w) b(v, w) \sum_{j=1}^\infty C_j(t) \delta(w - x_j) dw dv \\
&\quad + \int_{x_{i-1}}^{x_i} L(v, x_i) \int_v^\infty k(w) b(v, w) \sum_{j=1}^\infty C_j(t) \delta(w - x_j) dw dv \\
&= \sum_{j=i}^\infty k_j C_j(t) \int_{x_i}^{x_{i+1}} K(v, x_i) b(v, x_j) dv + \sum_{j=i}^\infty k_j C_j(t) \int_{x_{i-1}}^{x_i} L(v, x_i) b(v, x_j) dv \\
&= \sum_{j=i}^\infty k_j C_j(t) \left[\int_{x_i}^{x_{i+1}} K(v, x_i) b(v, x_j) dv + \int_{x_{i-1}}^{x_i} L(v, x_i) b(v, x_j) dv \right]
\end{aligned} \tag{A.9}$$

Using the relationship given by Eqs. (A.4) – (A.5), B now becomes:

$$B = \sum_{j=i}^\infty n_{ij} k_j C_j(t) \tag{A.10}$$

$$\begin{aligned}
n_{ij} &= \int_{x_i}^{x_{i+1}} \left(\frac{f_1(x_{i+1}) f_2(v) - f_1(v) f_2(x_{i+1})}{f_1(x_{i+1}) f_2(x_i) - f_1(x_i) f_2(x_{i+1})} \right) b(v, x_j) dv \\
&\quad + \int_{x_{i-1}}^{x_i} \left(\frac{f_1(x_{i-1}) f_2(v) - f_1(v) f_2(x_{i-1})}{f_1(x_{i-1}) f_2(x_i) - f_1(x_i) f_2(x_{i-1})} \right) b(v, x_j) dv
\end{aligned} \tag{A.11}$$

Given that the two integral properties of interest are $f_1(v) = v^\pi$ and $f_2(v) = v^\zeta$, n_{ij}

becomes:

$$n_{ij} = \int_{x_i}^{x_{i+1}} \left(\frac{x_{i+1}^\pi v^\zeta - v^\pi x_{i+1}^\zeta}{x_{i+1}^\pi x_i^\zeta - x_i^\pi x_{i+1}^\zeta} \right) b(v, x_j) dv + \int_{x_{i-1}}^{x_i} \left(\frac{x_{i-1}^\pi v^\zeta - v^\pi x_{i-1}^\zeta}{x_{i-1}^\pi x_i^\zeta - x_i^\pi x_{i-1}^\zeta} \right) b(v, x_j) dv \quad (\text{A.12})$$

For the conservation of the zeroth and the first moments, i.e. $\pi = 0$, $\zeta = 1$, n_{ij} collapse to:

$$n_{ij} = \int_{x_i}^{x_{i+1}} \left(\frac{x_{i+1} - v}{x_{i+1} - x_i} \right) b(v, x_j) dv + \int_{x_{i-1}}^{x_i} \left(\frac{v - x_{i-1}}{x_i - x_{i-1}} \right) b(v, x_j) dv \quad (\text{A.13})$$

Thus, the FP discretization scheme which conserves the zeroth and the first moments for general breakage is summarized as:

$$\frac{dC_i(t)}{dt} = \sum_{j=i}^{\infty} \left[\int_{x_i}^{x_{i+1}} \left(\frac{x_{i+1} - v}{x_{i+1} - x_i} \right) b(v, x_j) dv + \int_{x_{i-1}}^{x_i} \left(\frac{v - x_{i-1}}{x_i - x_{i-1}} \right) b(v, x_j) dv \right] k_j C_j(t) - k_i C_i(t) \quad (\text{A.14})$$

where the first integral term vanishes for $j = i$ (polymers with size x_i can only break into smaller sizes) and the second integral term vanishes for $i = 1$ (the smallest polymer size is the monomer).

A.2 MATLAB Code for The Simulation of Chain-End Scission using The Fixed Pivot Technique

```

%% Meshing
Mn_avg=4100; %number-average DP
Mw_avg=5430; %weight-average DP
xs=1; %monomer DP
xl=22496; %largest polymer DP
omega=Mn_avg/(Mw_avg-Mn_avg);
kappa=Mw_avg-Mn_avg;
Mtot=10; %initial starch concentration (g/L)

p=100; %pivots in discrete region
xgrid=1:l:p+1; xgrid=xgrid(:); %discrete region, x(1) to x(p+1)

N_con=600-(p+1); %pivots in continuous region
ratio=exp((1/N_con)*log((xl/xgrid(end)))); %geometric ratio

%geometric meshing
i=0; i=1:N_con+1;
xgrid_g=xgrid(end).*ratio.^(i-1);
xgrid=[xgrid(1:end-1);xgrid_g'];
xgrid(end)=xl;

Nb=length(xgrid); %number of pivot points
x_avg=zeros(Nb-1,1); i=1:Nb-1; x_avg(i,1)=(xgrid(i)+xgrid(i+1))/2;
x_avg=[xgrid(1);x_avg;xgrid(end)]; %v(i)

%% Initial MWD, Rate and Stoichiometric Kernels

%intermediate to calculate mass_tot
PP=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
(kappa.*gamma(omega)).*(x.*(162+18)-(xs.*(162+18)));

%molar concentration density (mol/L/DP)
mass_tot=quad(PP,x_avg(1),x_avg(end));
P=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
(kappa.*gamma(omega))).*Mtot./mass_tot;

%molar concentration (mol/L)
mole=zeros(Nb,1);
for i=1:Nb
    mole(i,1)=quad(P,x_avg(i),x_avg(i+1));
end

%rate kernel
kp=1/80;
S=kp*(xgrid).^1;

%n(i,k)
n=zeros(length(xgrid),length(xgrid));
n(1,1)=0; vm=1;
for k=2:length(xgrid)
    if (xgrid(k)-vm)<xgrid(2) && (xgrid(k)-vm)>xgrid(1)

```

```

        n(1,k)=(xgrid(2)-(xgrid(k)-vm)/(xgrid(2)-xgrid(1)));
    end
end
if sum(n(1,:))==0
    n(1,2)=1;
else
    disp('Oops.. Something is wrong with the n matrix!')
end

for i=2:length(xgrid)
    if (xgrid(i)-vm)<xgrid(i) && (xgrid(i)-vm)>xgrid(i-1)
        n(i,i)=(xgrid(i)-vm)-xgrid(i-1)/(xgrid(i)-xgrid(i-1));
    end
    for k=i+1:length(xgrid)
        if (xgrid(k)-vm)>xgrid(i) && (xgrid(k)-vm)<xgrid(i+1)
            n(i,k)=(xgrid(i+1)-(xgrid(k)-vm)/(xgrid(i+1)-xgrid(i)));
        end
        if (xgrid(k)-vm)>xgrid(i-1) && (xgrid(k)-vm)<xgrid(i)
            n(i,k)=(xgrid(k)-vm)-xgrid(i-1)/(xgrid(i)-xgrid(i-1));
        end
    end
    if sum(n(i,:))==0
        n(i,i+1)=1;
    end
end

%% Run the Simulation
tstart=0; global tfinal; tfinal=1000;
tgrid(:,1)=tstart:tfinal; %simulation time

options=odeset('NonNegative',1:Nb,'Stats','on',...
    'Refine',8,'AbsTol',1e-16,'RelTol',1e-6);

[t,y]=ode45(@(t,y) chain_end_fp_ode(t,y,S,Nb,n,p),tgrid,mole, \
options);

```

```

function dCdt=chain_end_fp_ode(t,y,S,Ng,n,p)

%State variables
C=zeros(Ng,1);
for i=1:Ng
    C(i)=y(i);
end

%ODEs
dC=zeros(1,Ng);

dC(1)=2*S(2)*C(2)+sum(S(3:Ng).*C(3:Ng));

for i=2:p
    dC(i)=S(i+1)*C(i+1)-S(i)*C(i);
end

dC(p+1)=n(p+1,p+2)*S(p+2)*C(p+2)-S(p+1)*C(p+1);

for i=p+2:Ng-1
    dC(i)=n(i,i)*S(i)*C(i)+n(i,i+1)*S(i+1)*C(i+1)-S(i)*C(i);
end

dC(Ng)=n(Ng,Ng)*S(Ng)*C(Ng)-S(Ng)*C(Ng);

%Output the state variables
dCdt=dC';

```

A.3 MATLAB Code for The Simulation of Chain-End Scission using The Exact Solution

```

%% Meshing
Mn_avg=4100; %number-average DP
Mw_avg=5430; %weight-average DP
xs=1; %monomer DP
xl=22496; %largest polymer DP
omega=Mn_avg/(Mw_avg-Mn_avg);
kappa=Mw_avg-Mn_avg;
Mtot=10; %initial starch concentration (g/L)

xgridl=xs:1:xl; xgridl=xgridl(:); %arithmetic meshing

Nbl=length(xgridl); %number of pivot points
x_avgl=zeros(Nbl-1,1); i=1:Nbl-1; x_avgl(i,1)=(xgridl(i)+xgridl(i+1)) /
/2;
x_avgl=[xgridl(1);x_avgl;xgridl(end)]; %v(i)

%% Initial MWD, Rate and Stoichiometric Kernels

%intermediate to calculate mass_tot
PP=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
(kappa.*gamma(omega))).*(x.*(162+18)-(xs.*(162+18)));

%Molar concentration density (mol/L/DP)
mass_tot=quad(PP,x_avgl(1),x_avgl(end));
P=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
(kappa.*gamma(omega))).*Mtot./mass_tot;

%Molar concentration (mol/L)
mole=zeros(Nbl,1);
for i=1:Nbl
    mole(i,1)=quad(P,x_avgl(i),x_avgl(i+1));
end

%rate kernel
kp=1/80;
S=kp*(xgridl).^1;

%% Run the Simulation
tstart=0; global tfinal; tfinal=1000;
tgrid(:,1)=tstart:tfinal; %simulation time

options=odeset('NonNegative',1:Nbl,'Stats','on',...
'Refine',8,'AbsTol',1e-16,'RelTol',1e-6);

[t1,y1] = ode45(@(t,y) chain_end_ode(t,y,S,Nbl),tgrid,mole,options);

```

```

function dCdt=chain_end_ode(t,y,S,Nb)

%State variables
C=zeros(Nb,1);
for i=1:Nb
    C(i,1)=y(i);
end

%ODEs
dC=zeros(1,Nb);

dC(1)=2*S(2)*C(2)+sum(S(3:Nb).*C(3:Nb));

for i=2:Nb-1
    dC(i)=S(i+1)*C(i+1)-S(i)*C(i);
end

dC(Nb)=-S(Nb)*C(Nb);

%Output the state variables
dCdt=dC';

```

A.4 MATLAB Code for The Simulation of Chain-End Scission on Stickel's

Example (Stickel and Griggs,2012) using The Fixed Pivot Technique

```
%% Meshing
xl=827; %largest polymer DP
omega=400;
kappa=1.5;
Cin=100;

xgrid=1:1:10; xgridg=1:1.5:827; xgrid=[xgrid xgridg];
xgrid=xgrid(:); %arithmetic progression

Nb=length(xgrid); %number of pivot points
x_avg=zeros(Nb-1,1); i=1:Nb-1; x_avg(i,1)=(xgrid(i)+xgrid(i+1))/2;
x_avg=[xgrid(1);x_avg;xgrid(end)]; %v(i)

%% Initial MWD
P=zeros(Nb,1);
for jj=1:length(xgrid)
    intermediat1=(xgrid(jj)/kappa)*exp(-xgrid(jj)/kappa)/(omega-1);
    for ii=1:omega-2
        intermediat1=intermediat1*(xgrid(jj)/kappa)/ii;
    end
    P(jj,1)=(Cin*intermediat1)/(kappa);
end
mole_mwd=spline(xgrid,P);

% molar concentration
mole=zeros(Nb,1);
for i=1:Nb
    mole(i,1)=quad(@(x) ppval(mole_mwd,x),x_avg(i),x_avg(i+1));
    if mole(i,1)<0
        mole(i,1)=mole(i,1)*-1;
    end
    if abs(mole(i,1))<1e-16
        mole(i,1)=0;
    end
end

%% Rate and Stoichiometric Kernels

%rate kernel
kp=1;
S=kp*(xgrid).^0;

%n(i,k)
n=zeros(length(xgrid),length(xgrid));
n(1,1)=0; vm=1;
for k=2:length(xgrid)
    if (xgrid(k)-vm)<xgrid(2) && (xgrid(k)-vm)>xgrid(1)
        n(1,k)=(xgrid(2)-(xgrid(k)-vm))/(xgrid(2)-xgrid(1));
    end
end
end
for i=2:length(xgrid)
```

```

        if (xgrid(i)-vm)<xgrid(i) && (xgrid(i)-vm)>xgrid(i-1)
            n(i,i)=((xgrid(i)-vm)-xgrid(i-1))/(xgrid(i)-xgrid(i-1));
        end
        for k=i+1:length(xgrid)
            if (xgrid(k)-vm)>xgrid(i) && (xgrid(k)-vm)<xgrid(i+1)
                n(i,k)=(xgrid(i+1)-(xgrid(k)-vm))/(xgrid(i+1)-xgrid(i));
            end
            if (xgrid(k)-vm)>xgrid(i-1) && (xgrid(k)-vm)<xgrid(i)
                n(i,k)=((xgrid(k)-vm)-xgrid(i-1))/(xgrid(i)-xgrid(i-1));
            end
        end
    end
end

%% Run the Simulation
tstart=0; global tfinal; tfinal=600;
tgrid(:,1)=tstart:tfinal; %simulation time
options=odeset('NonNegative',1:Nb,'Stats','on',...
    'Refine',8,'AbsTol',1e-16,'RelTol',1e-6);
[t,y] = ode45(@chain_end_fp_ode(t,y,S,Nb,n,total),tgrid,...
    mole,options);

function dCdt=chain_end_fp_ode(t,y,S,Ng,n,p)

%State variables
C=zeros(Ng,1);
for i=1:Ng
    C(i)=y(i);
end

%ODEs
dC=zeros(1,Ng);

dC(1)=2*S(2)*C(2)+sum(S(3:Ng).*C(3:Ng));

for i=2:p
    dC(i)=S(i+1)*C(i+1)-S(i)*C(i);
end

dC(p+1)=n(p+1,p+2)*S(p+2)*C(p+2)-S(p+1)*C(p+1);

for i=p+2:Ng-1
    dC(i)=n(i,i)*S(i)*C(i)+n(i,i+1)*S(i+1)*C(i+1)-S(i)*C(i);
end

dC(Ng)=n(Ng,Ng)*S(Ng)*C(Ng)-S(Ng)*C(Ng);

%Output the state variables
dCdt=dC';

```


A.5 MATLAB Code for The Simulation of Chain-End Scission on Stickel's

Example (Stickel and Griggs, 2012) using The Exact Solution

```
%% Meshing
xl=827; %largest polymer DP
omega=400;
kappa=1.5;
Cin=100;

xgridl=1:1:xl; xgridl=xgridl(:); %arithmtetic meshing

Nbl=length(xgridl); %number of pivot points
x_avgl=zeros(Nbl-1,1); i=1:Nbl-1; x_avgl(i,1)=(xgridl(i)+xgridl(i+1))/2;
x_avgl=[xgridl(1);x_avgl;xgridl(end)]; %v(i)

%% Initial MWD
P=zeros(Nbl,1);
for jj=1:length(xgridl)
    intermediatel=(xgridl(jj)/kappa)*exp(-xgridl(jj)/kappa)/(omega-1);
    for ii=1:omega-2
        intermediatel=intermediatel*(xgridl(jj)/kappa)/ii;
    end
    P(jj,1)=(Cin*intermediatel)/(kappa);
end
mole_mwdl=spline(xgridl,P);

% molar concentration
molel=zeros(Nbl,1);
for i=1:Nbl
    molel(i,1)=quad(@(x) ppval(mole_mwdl,x),x_avgl(i),x_avgl(i+1));
    if molel(i,1)<0
        molel(i,1)=molel(i,1)*-1;
    end
    if abs(molel(i,1))<1e-16
        molel(i,1)=0;
    end
end

%% Rate Kernel
kp=1;
S=kp*(xgridl).^0;

%% Run the Simulation
tstart=0; global tfinal;
tgrid(:,1)=tstart:tfinal; %simulation time
options=odeset('NonNegative',1:Nbl,'Stats','on',...
    'Refine',8,'AbsTol',1e-16,'RelTol',1e-6);
[t,y1] = ode45(@(t,y) chain_end_ode(t,y,S,Nbl),tgrid,molel,options);
```



```

function dCdt=chain_end_ode(t,y,S,Nb)

%State variables
C=zeros(Nb,1);
for i=1:Nb
    C(i,1)=y(i);
end

%ODEs
dC=zeros(1,Nb);

dC(1)=2*S(2)*C(2)+sum(S(3:Nb).*C(3:Nb));

for i=2:Nb-1
    dC(i)=S(i+1)*C(i+1)-S(i)*C(i);
end

dC(Nb)=-S(Nb)*C(Nb);

%Output the state variables
dCdt=dC';

```

A.6 MATLAB Code for Calculating p_{min}

```
N=22496; %maximum DP
vm=1; %vm

total_pivots=600; %p+q

p=0;
n_pivots=0;
while n_pivots < total_pivots
    p=p+1;
    n_pivots=(p+1)+log(N/(p+vm))/log((p+2*vm)/(p+vm));
end

fprintf('The value of p = %d at a total number of pivots = %d.\n',...
    p,total_pivots)
```

APPENDIX B: MODELLING RANDOM SCISSION

B.1 Deriving The Fixed Pivot Equations for Discrete Random Scission

The fully discrete (or exact) PBEs for pure random scission are given by Eqs. (4.38) - (4.40) in Section 4.2.1. In the work of S. Kumar and Ramkrishna (1996a), the relevant kernel necessary for the approximation of discrete random scission using the FP technique had been identified. To fill in the missing details in the original paper (S. Kumar & Ramkrishna, 1996a), this appendix derives the FP equations for random scission using the reported kernel. According to their work, the stoichiometric kernel for random scission is:

$$b^{\alpha}(v, w) = \frac{2\delta(v - iv_m)}{w - 1} \quad i = 1, 2, 3, \dots \quad (\text{B.1})$$

where $v_m = 1$ is the DP of the monomer. Substituting this into Eq. (A.14) in Appendix A, the FP equation for random scission is:

$$\frac{dC_i(t)}{dt} = \sum_{j=i}^{p+q} n_{ij}^{\alpha} k_j^{\alpha} C_j(t) - k_i^{\alpha} C_i(t) \quad (\text{B.2})$$

$$n_{ij}^{\alpha} = \int_{x_i}^{x_{i+1}} \left(\frac{x_{i+1} - v}{x_{i+1} - x_i} \right) \left[\frac{2\delta(v - i)}{x_j - 1} \right] dv + \int_{x_{i-1}}^{x_i} \left(\frac{v - x_{i-1}}{x_i - x_{i-1}} \right) \left[\frac{2\delta(v - i)}{x_j - 1} \right] dv \quad (\text{B.3})$$

The following evaluates $dC_i(t)/dt$ for different i 's.

When $i = 1$, the death term as well as the second integral of $n_{1,j}^{\alpha}$ is zero. The remaining terms in $n_{1,j}^{\alpha}$ can be evaluated using the limiting process presented in Section 4.1.3 as:

$$\begin{aligned}
n_{1,j}^\alpha &= \lim_{\chi \rightarrow 0} \int_{x_1}^{x_2} \left(\frac{x_2 - v}{x_2 - x_1} \right) \left(\frac{2}{x_j - 1} \right) \delta[v - (1 + \chi)] dv \\
&= \lim_{\chi \rightarrow 0} \left\{ \left[\frac{x_2 - (1 + \chi)}{x_2 - x_1} \right] \left[\frac{2}{x_j - 1} \right] \right\} \\
&= \left[\frac{x_2 - 1}{x_2 - x_1} \right] \left[\frac{2}{x_j - 1} \right] \\
&= \frac{2}{x_j - 1}
\end{aligned} \tag{B.4}$$

Since $x_1 = 1$, for $i = 1$, the resulting FP equation is:

$$\frac{dC_1(t)}{dt} = \sum_{j=2}^{p+q} \left[\frac{2}{x_j - 1} \right] k_j^\alpha C_j(t) \tag{B.5}$$

For $i = 2$ to p (i.e. the discrete region), at $j = i$, the first integral vanishes because polymers do not grow in size during a pure depolymerization phenomena. Moreover, because of the limiting process, the second integral of n_{ij}^α will always be zero, and thus the following equation is obtained:

$$n_{ij}^\alpha = \begin{cases} 0 & j = i \\ \left(\frac{x_{i+1} - i}{x_{i+1} - x_i} \right) \left[\frac{2}{x_j - 1} \right] & j > i \end{cases} \tag{B.6}$$

Again, since $x_i = i$ in the discrete region, the resulting FP equation is:

$$\frac{dC_i(t)}{dt} = \sum_{j=i+1}^{p+q} \left[\frac{2}{x_j - 1} \right] k_j^\alpha C_j(t) - k_i^\alpha C_i(t) \quad i = 2, 3, \dots, p \tag{B.7}$$

Next, moving to $i = p+1$, x_{p+1} is situated at the border between the discrete and the continuous region. Because of this, it is reasonable to assume that a transition from the discontinuous to a fully continuous kernel is involved. Therefore, $n_{p+1,j}^\alpha$ assumes the following form:

$$n_{p+1,j}^\alpha = \int_{x_{p+1}}^{x_{p+2}} \left(\frac{x_{p+2} - v}{x_{p+2} - x_{p+1}} \right) \left[\frac{2}{x_j - 1} \right] dv + \int_{x_p}^{x_{p+1}} \left(\frac{v - x_p}{x_{p+1} - x_p} \right) \left[\frac{2\delta(v - [p+1])}{x_j - 1} \right] dv \quad (\text{B.8})$$

where to the right of x_{p+1} , a continuous random scission kernel is used whereas the discontinuous Dirac delta kernel is retained to the left of x_{p+1} . Similarly, at $j = p+1$, the first integral vanishes while the second integral is always zero using the limiting process.

This results in:

$$n_{p+1,j}^\alpha = \begin{cases} 0 & j = p+1 \\ \frac{x_{p+2} - x_{p+1}}{x_j - 1} & j > p+1 \end{cases} \quad (\text{B.9})$$

The corresponding FP equation is then given as:

$$\frac{dC_{p+1}(t)}{dt} = \sum_{j=p+2}^{p+q} \left[\frac{x_{p+2} - x_{p+1}}{x_j - 1} \right] k_j^\alpha C_j(t) - k_{p+1}^\alpha C_{p+1}(t) \quad (\text{B.10})$$

Finally, at $i = p+2$ to $p+q$, the continuum approximation fully holds, and thus n_{ij}^α can be written as:

$$n_{ij}^\alpha = \int_{x_i}^{x_{i+1}} \left(\frac{x_{i+1} - v}{x_{i+1} - x_i} \right) \left[\frac{2}{x_j - 1} \right] dv + \int_{x_{i-1}}^{x_i} \left(\frac{v - x_{i-1}}{x_i - x_{i-1}} \right) \left[\frac{2}{x_j - 1} \right] dv \quad (\text{B.11})$$

Evaluation of this using similar rationale as presented above results in:

$$n_{ij}^\alpha = \begin{cases} \frac{x_i - x_{i-1}}{x_j - 1} & j = i \\ \frac{x_{i+1} - x_i}{x_j - 1} + \frac{x_i - x_{i-1}}{x_j - 1} & j > i \end{cases} \quad (\text{B.12})$$

The FP equation (for $i = p+2$ to $p+q$) is thus given as:

$$\frac{dC_i(t)}{dt} = \left[\frac{x_i - x_{i-1}}{x_i - 1} \right] k_i^\alpha C_i(t) + \sum_{j=i+1}^{p+q} \left[\frac{x_{i+1} - x_i}{x_j - 1} + \frac{x_i - x_{i-1}}{x_j - 1} \right] k_j^\alpha C_j(t) - k_i^\alpha C_i(t) \quad (\text{B.13})$$

Equations (B.5), (B.7), (B.10), and (B.13) thus represent the FP equations for random scission, and they form the basis for the discussion presented in Section 4.2.3.

B.2 Simplifying The Fixed Pivot Expression for Random Scission

From the FP equation for pure random scission [Eq. (4.59), Section 4.2.3], the rate of change in the concentration of the polymer at the i -th pivot can be re-written as follows:

$$\frac{dC_i(t)}{dt} = \sum_{j=i+1}^{p+q} \left[\frac{x_{i+1}-x_i}{x_j-1} + \frac{x_i-x_{i-1}}{x_j-1} \right] k_j^\alpha C_j(t) - \left[\frac{x_{i-1}-1}{x_i-1} \right] k_i^\alpha C_i(t) \quad (\text{B.14})$$

Expanding the summation term, and with $r = \frac{x_i - x_{i-1}}{x_{i-1} - x_{i-2}} = \frac{v_i - v_{i-1}}{v_{i-1} - v_{i-2}}$, the above

expression can be written in terms of c_i (the average molar concentration density) as shown below:

$$\begin{aligned} \frac{dC_i(t)}{dt} = & \left[\frac{x_{i+1}-x_i}{x_{i+1}-1} + \frac{x_i-x_{i-1}}{x_{i+1}-1} \right] k_{i+1}^\alpha c_{i+1}(t) r (v_{i+1} - v_i) \\ & + \left[\frac{x_{i+1}-x_i}{x_{i+2}-1} + \frac{x_i-x_{i-1}}{x_{i+2}-1} \right] k_{i+2}^\alpha c_{i+2}(t) r^2 (v_{i+1} - v_i) + \dots \\ & + \left[\frac{x_{i+1}-x_i}{x_{p+q}-1} + \frac{x_i-x_{i-1}}{x_{p+q}-1} \right] k_{p+q}^\alpha c_{p+q}(t) r^{p+q-i} (v_{i+1} - v_i) \\ & - \left[\frac{x_{i-1}-1}{x_i-1} \right] k_i^\alpha c_i(t) (v_{i+1} - v_i) \end{aligned} \quad (\text{B.15})$$

Dividing both sides of the equation with $(v_{i+1} - v_i)$, the following equation is obtained:

$$\begin{aligned} \frac{dc_i(t)}{dt} = & r \left[\frac{x_{i+1}-x_i}{x_{i+1}-1} + \frac{x_i-x_{i-1}}{x_{i+1}-1} \right] k_{i+1}^\alpha c_{i+1}(t) + r^2 \left[\frac{x_{i+1}-x_i}{x_{i+2}-1} + \frac{x_i-x_{i-1}}{x_{i+2}-1} \right] k_{i+2}^\alpha c_{i+2}(t) + \dots \\ & + r^{p+q-i} \left[\frac{x_{i+1}-x_i}{x_{p+q}-1} + \frac{x_i-x_{i-1}}{x_{p+q}-1} \right] k_{p+q}^\alpha c_{p+q}(t) - \left[\frac{x_{i-1}-1}{x_i-1} \right] k_i^\alpha c_i(t) \end{aligned} \quad (\text{B.16})$$

Equation (B.16) can be further simplified by using the relationship

$x_i - x_{i-1} = (x_{i+1} - x_i)/r$ to yield:

$$\begin{aligned}
\frac{dc_i(t)}{dt} &= (x_{i+1} - x_i) \left\{ \left[\frac{(r+1)}{x_{i+1}-1} \right] k_{i+1}^\alpha c_{i+1}(t) + r \left[\frac{(r+1)}{x_{i+2}-1} \right] k_{i+2}^\alpha c_{i+2}(t) + \dots \right. \\
&\quad \left. + r^{p+q-i-1} \left[\frac{(r+1)}{x_{p+q}-1} \right] k_{p+q}^\alpha c_{p+q}(t) \right\} - \left[\frac{x_{i-1}-1}{x_i-1} \right] k_i^\alpha c_i(t) \\
&= (x_{i+1} - x_i) \sum_{j=i+1}^{p+q} r^{j-(i+1)} \left[\frac{(r+1)}{x_j-1} \right] k_j^\alpha c_j(t) - \left[\frac{x_{i-1}-1}{x_i-1} \right] k_i^\alpha c_i(t)
\end{aligned} \tag{B.17}$$

Because Eq. (B.17) cannot be further simplified by solely using the expression $r = (x_{i+1} - x_i) / (x_i - x_{i-1})$, the expression $x_{i+1} = rx_i$ for a geometric mesh has to be invoked. Therefore, $x_{i+1} = r^{i+1-(p+1)} x_{p+1}$, $x_i = r^{i-(p+1)} x_{p+1}$ and $x_{i-1} = r^{i-1-(p+1)} x_{p+1}$.

Substituting these in Eq. (B.17) and further simplifying results in:

$$\frac{dc_i(t)}{dt} = x_{p+1} \sum_{j=i+1}^{p+q} \left[\frac{r^{j-p} - r^{j-p-2}}{x_j - 1} \right] k_j^\alpha c_j(t) - \left[\frac{r^{i-1-(p+1)} x_{p+1} - 1}{r^{i-(p+1)} x_{p+1} - 1} \right] k_i^\alpha c_i(t) \tag{B.18}$$

This is Eq. (4.69) in Section 4.3.

B.3 Proof that the Fixed Pivot Technique Over-Predicts for Random Scission

To facilitate the following proof, the FP equation for random scission presented in Section 4.3 is reproduced here below with the $m \geq 2$ terms written in full:

$$\frac{dc_i(t)}{dt} = \sum_{j=i+1}^{p+q} f(\varepsilon) k_j^\alpha c_j(t) - g(\varepsilon) k_i^\alpha c_i(t) \tag{B.19}$$

$$f(\varepsilon) = \left[\frac{u^{-1}}{x_j - 1} \right] \left\{ 2(u + \varepsilon) + \sum_{m=2}^{\infty} \left[\frac{(j-p-2)! [2(j-p) - (m+1)]}{(j-p-m)! (m-1)!} \right] (u + \varepsilon)^m \right\} \tag{B.20}$$

$$g(\varepsilon) = \frac{p + u^{-1} (i-p-2) (u + \varepsilon) + u^{-1} \sum_{m=2}^{\infty} \frac{(i-p-2)!}{(i-p-2-m)! m!} (u + \varepsilon)^m}{p + u^{-1} (i-p-1) (u + \varepsilon) + u^{-1} \sum_{m=2}^{\infty} \frac{(i-p-1) (i-p-2)!}{(i-p-1-m)! (i-p-2-m)! m!} (u + \varepsilon)^m} \tag{B.21}$$

where $u = (1 + p)^{-1}$.

To prove that $g(\varepsilon)$ always decreases further from unity, the expression can be re-written as:

$$g(\varepsilon) = \frac{p + u^{-1} \sum_{m=1}^{\infty} \Omega_m (u + \varepsilon)^m}{p + u^{-1} \sum_{m=1}^{\infty} \frac{(i - p - 1)}{(i - p - 1 - m)} \Omega_m (u + \varepsilon)^m}; \quad \Omega_m = \frac{(i - p - 2)!}{(i - p - 2 - m)! m!} \quad (\text{B.22})$$

Since $\frac{(i - p - 1)}{(i - p - 1 - m)} > 1$, the denominator of $g(\varepsilon)$ is always larger than the numerator,

and thus $g(\varepsilon)$ decreases further from unity as ε increases.

To prove that $f(\varepsilon)$ increases further from $2/(x_j - 1)$ as r deviates from the limit of arithmetic mesh, the proof is deduced from the following expressions:

$$f(\varepsilon) = 2 \left[\frac{u^{-1}}{x_j - 1} \right] (u + \varepsilon) + \left[\frac{u^{-1}}{x_j - 1} \right] \sum_{m=2}^{\infty} [2(j - p) - (m + 1)] \tilde{\Omega}_m; \quad (\text{B.23})$$

$$\tilde{\Omega}_m = \left[\frac{(j - p - 2)!}{(j - p - m)!(m - 1)!} \right] (u + \varepsilon)^m$$

In Eq. (B.23), the first term on the RHS of $f(\varepsilon)$ is always greater or equal to $2/(x_j - 1)$ because $\varepsilon \geq 0$ (due to the constraint $r \geq 1 + u$). Thus, as long as the second term remains positive, $f(\varepsilon) > 2/(x_j - 1)$. Therefore, the task here is to show that $2(j - p) - (m + 1)$ in the second term is always positive as $\tilde{\Omega}_m > 0$. From Eq. (4.75) in Section 4.3, binomial expansion of the second term on the RHS requires that the value of $m \leq j - p - 2$. In addition, the most narrow distance between j and p occurs at $j = p + 2$. With these, it follows that:

$$\begin{aligned}
2(j-p)-(m+1) &= 2(j-p)-(j-p-2+1) \\
&= j-p+1 \\
&= (p+2)-p+1 \\
&= 3(>0)
\end{aligned}
\tag{B.24}$$

Hence, it is shown that the second term on the RHS of Eq. (B.23) is always positive, and as ε increases, $f(\varepsilon)$ departs further from $2/(x_j-1)$.

B.4 MATLAB Code for The Simulation of Random Scission on Stickel's

Example (Stickel and Griggs, 2012) Using The Fixed Pivot Technique

```
%% Meshing
xl=827; %largest polymer DP
omega=400;
kappa=1.5;
Cin=100;

%% Arithmetic meshing
p=10;
xgrid=1:1:p; xgridg=p+1:1.5:827; xgrid=[xgrid xgridg];
xgrid=xgrid(:);

Nb=length(xgrid); %number of pivot points
x_avg=zeros(Nb-1,1); i=1:Nb-1; x_avg(i,1)=(xgrid(i)+xgrid(i+1))/2;
x_avg=[xgrid(1);x_avg;xgrid(end)]; %v(i)

%% Initial MWD
P=zeros(Nb,1);
for jj=1:length(xgrid)
    intermediatel=(xgrid(jj)/kappa)*exp(-xgrid(jj)/kappa)/(omega-1);
    for ii=1:omega-2
        intermediatel=intermediatel*(xgrid(jj)/kappa)/ii;
    end
    P(jj,1)=(Cin*intermediatel)/(kappa);
end
mole_mwd=spline(xgrid,P);

%molar concentration
mole=zeros(Nb,1);
for i=1:Nb
    mole(i,1)=quad(@(x) ppval(mole_mwd,x),x_avg(i),x_avg(i+1));
    if mole(i,1)<0
        mole(i,1)=mole(i,1)*-1;
    end
    if abs(mole(i,1))<1e-16
        mole(i,1)=0;
    end
end

%% Rate Kernel for Random Scission
kp_a=1;
S_a=zeros(Nb,1);
for i=2:Nb
    S_a(i)=kp_a*(xgrid(i)-1)^0;
end

%% n(i,k) - Random Scission
xgridm=xgrid;
nn=zeros(length(xgridm),length(xgridm));
nn(1,1)=0;
i=1;
for k=2:length(xgridm)
```

```

        nn(1,k)=2/(xgridm(k)-1);
    end
    for i=2:p
        for k=i+1:length(xgridm)
            nn(i,k)=2/(xgridm(k)-1);
        end
    end

    for k=i+1:length(xgridm)
        nn(p+1,k)=(xgridm(p+2)-xgridm(p+1))/(xgridm(k)-1);
    end

    for i=p+2:length(xgridm)
        nn(i,i)=(xgridm(i)-xgridm(i-1))/(xgridm(i)-1);
        for k=i+1:length(xgridm)
            nn(i,k)=(xgridm(i+1)-xgridm(i))/(xgridm(k)-1)+...
                (xgridm(i)-xgridm(i-1))/(xgridm(k)-1);
        end
    end
end

%% Run the Simulation
tstart=0; global tfinal; tfinal=60;
tgridl(:,1)=tstart:60/100:tfinal; %simulation time

options=odeset('NonNegative',1:Nb,'Stats','on',...
    'Refine',8,'AbsTol',1e-22,'RelTol',1e-6);

[t,y] = ode15s(@(t,y) random_fp_ode(t,y,S_a,Nb,nn,p),...
    tgridl,mole,options);

```

```

function dCdt=random_fp_ode(t,y,S_a,Ng,nn,p)

%State variables
C=zeros(Ng,1);
for i=1:Ng
    C(i)=y(i);
end

%ODEs
dC=zeros(1,Ng);

summation=0;
for j=2:Ng
    summation=summation+S_a(j)*nn(1,j)*C(j);
end

dC(1)=summation;

for i=2:p
    summation=0;
    for j=i+1:Ng
        summation=summation+S_a(j)*nn(i,j)*C(j);
    end
    dC(i)=summation-S_a(i)*C(i);
end

summation=0;
for j=p+2:Ng
    summation=summation+S_a(j)*nn(p+1,j)*C(j);
end
dC(p+1)=summation-S_a(p+1)*C(p+1);

for i=p+2:Ng-1
    summation=0;
    for j=i:Ng
        summation=summation+S_a(j)*nn(i,j)*C(j);
    end
    dC(i)=summation-S_a(i)*C(i);
end

dC(Ng)=S_a(Ng)*nn(Ng,Ng)*C(Ng)-S_a(Ng)*C(Ng);

%Output the state variables
dCdt=dC';

```

B.5 MATLAB Code for The Simulation of Random Scission on Stickel's

Example (Stickel and Griggs, 2012) Using The Exact Solution

```
%% Meshing
xl=827; %largest polymer DP
omega=400;
kappa=1.5;
Cin=100;

xgrid1=1:1:xl; xgrid1=xgrid1(:); %arithmtetic progression

Nbl=length(xgrid1); %number of pivot points
x_avgl=zeros(Nbl-1,1); i=1:Nbl-1; x_avgl(i,1)=(xgrid1(i)+xgrid1(i+1))/2;
x_avgl=[xgrid1(1);x_avgl;xgrid1(end)]; %v(i)

%% Initial MWD
P=zeros(Nbl,1);
for jj=1:length(xgrid1)
    intermediatel=(xgrid1(jj)/kappa)*exp(-xgrid1(jj)/kappa)/(omega-1);
    for ii=1:omega-2
        intermediatel=intermediatel*(xgrid1(jj)/kappa)/ii;
    end
    P(jj,1)=(Cin*intermediatel)/(kappa);
end
mole_mwd1=spline(xgrid1,P);

% molar concentration
mole1=zeros(Nbl,1);
for i=1:Nbl
    mole1(i,1)=quad(@(x) ppval(mole_mwd1,x),x_avgl(i),x_avgl(i+1));
    if mole1(i,1)<0
        mole1(i,1)=mole1(i,1)*-1;
    end
    if abs(mole1(i,1))<1e-16
        mole1(i,1)=0;
    end
end

%% Rate Kernel for Random Scission
kp_a=1;
S_al=kp_a*(xgrid1-1).^0;

%% Coefficient matrix for Random Scission
n_full=zeros(Nbl,Nbl);
for i=1:Nbl
    for j=i+1:Nbl
        n_full(i,j)=(2/(j-1))*S_al(j);
    end
end

for i=2:Nbl
    n_full(i,i)=n_full(i,i)-S_al(i);
```

```

end

%% Break into series of matrices
msz=827;

mrt{1}=zeros(msz,msz);

for i=1:Nbl/msz
    i1=Nbl-msz*i+1;
    i2=Nbl-msz*(i-1);
    nrt{i}=n_full(i1:i2,i1:i2);
    if i>1
        mrt{i} = n_full(i1:i2,i2+1:end);
    end
end

%% Chunk by chunk solving
tstart=0; global tfinal; tfinal=60;
tgrid(:,1)=tstart:60/100:tfinal;

yn=mole1;
yy=[];
ttime=[];
stats_exact=[];

time_full_sim=toc;

mrt{1}=zeros(msz,msz);
for j=1:Nbl/msz
    global ode_count;
    ode_count=j;
    i1=Nbl-msz*j+1;
    i2=Nbl-msz*(j-1);

    options=odeset('NonNegative',1:length(yn(i1:i2)),'Stats','on',...
        'Refine',8,'AbsTol',1e-22,'RelTol',1e-6,...
        'JPattern',triu(ones(size(nrt{j},1),size(nrt{j},2))));
    sol = ode15s(@(t,y) chunking_ode2(t,y,nrt{j},mrt{j},ttime...
        ,yy,j,msz),tgrid,yn(i1:i2),options);

    stats_exact{j}=sol.stats;
    ttime{j} = sol.x';
    yy{j}=sol.y';
end

function ypr = chunking_ode2(t,y,nrt,mrt,ttime,yy,j,msz)

yyl = zeros(msz,1);

if j>1
    yy1=[];
    for ii=1:j-1
        yysol=interp1(ttime{ii},yy{ii},t,'linear');
        yy1=[yysol yy1];
    end
end

ypr = nrt*y(:)+mrt*yy1(:);

ypr=ypr(:);

```

APPENDIX C: MODELLING SIMULTANEOUS RANDOM AND CHAIN-END SCISSION

C.1 MATLAB Code for The Simulation of Simultaneous Random and Chain- End Scissions Using The Fixed Pivot Technique

```

%% Meshing
Mn_avg=4100; %number-average DP
Mw_avg=5430; %weight-average DP
xs=1; %monomer DP
xl=22496; %largest polymer DP
omega=Mn_avg/(Mw_avg-Mn_avg);
kappa=Mw_avg-Mn_avg;
Mtot=10; %initial starch concentration (g/L)

p=100; %%pivots in discrete region
xgrid=1:1:p+1; xgrid=xgrid(:); %discrete region, x(1) to x(p+1)

N_con=200-(p+1); %%pivots in continuous region
ratio=exp((1/N_con)*log((xl/xgrid(end)))); %geometric ratio

%geometric meshing
i=0; i=1:N_con+1;
xgrid_g=xgrid(end).*ratio.^(i-1);
xgrid=[xgrid(1:end-1);xgrid_g'];
xgrid(end)=xl;

Nb=length(xgrid); %number of pivot points
x_avg=zeros(Nb-1,1); i=1:Nb-1; x_avg(i,1)=(xgrid(i)+xgrid(i+1))/2;
x_avg=[xgrid(1);x_avg;xgrid(end)]; %v(i)

%% Initial MWD

%intermediate to calculate mass_tot
PP=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
(kappa.*gamma(omega)).*(x.*162+18)-(xs.*162+18));

%molar concentration density (mol/L/DP)
mass_tot=quad(PP,x_avg(1),x_avg(end));
P=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
(kappa.*gamma(omega)).*Mtot./mass_tot;

%molar concentration (mol/L)
mole=zeros(Nb,1);
for i=1:Nb
    mole(i,1)=quad(P,x_avg(i),x_avg(i+1));
end

%% Rate Kernel for Chain-End Scission
kp_g=1/80;
S_g=kp_g*(xgrid).^1;

%% Rate Kernel for Random Scission
kp_a=1/200;
S_a=zeros(Nb,1);
for i=2:Nb
    S_a(i)=kp_a*(xgrid(i)-1)^1;

```



```

end

%% n(i,k) for Chain-End Scission
n=zeros(length(xgrid),length(xgrid));
n(1,1)=0; vm=1;
for k=2:length(xgrid)
    if (xgrid(k)-vm)<xgrid(2) && (xgrid(k)-vm)>xgrid(1)
        n(1,k)=(xgrid(2)-(xgrid(k)-vm))/(xgrid(2)-xgrid(1));
    end
end
if sum(n(1,:))==0
    n(1,2)=1;
else
    disp('Oops.. Something is wrong with the n matrix!')
end

for i=2:length(xgrid)
    if (xgrid(i)-vm)<xgrid(i) && (xgrid(i)-vm)>xgrid(i-1)
        n(i,i)=(xgrid(i)-vm-xgrid(i-1))/(xgrid(i)-xgrid(i-1));
    end
    for k=i+1:length(xgrid)
        if (xgrid(k)-vm)>xgrid(i) && (xgrid(k)-vm)<xgrid(i+1)
            n(i,k)=(xgrid(i+1)-(xgrid(k)-vm))/(xgrid(i+1)-xgrid(i));
        end
        if (xgrid(k)-vm)>xgrid(i-1) && (xgrid(k)-vm)<xgrid(i)
            n(i,k)=(xgrid(k)-vm-xgrid(i-1))/(xgrid(i)-xgrid(i-1));
        end
    end
    if sum(n(i,:))==0
        n(i,i+1)=1;
    end
end

total=zeros(Nb,1); for i=1:Nb; total(i)=sum(n(:,i)); end;

%% n(i,k) for Random Scission
xgridm=xgrid;
nn=zeros(length(xgridm),length(xgridm));
nn(1,1)=0;
i=1;
for k=2:length(xgridm)
    nn(1,k)=2/(xgridm(k)-1);
end
for i=2:p
    for k=i+1:length(xgridm)
        nn(i,k)=2/(xgridm(k)-1);
    end
end

for k=i+1:length(xgridm)
    nn(p+1,k)=(xgridm(p+2)-xgridm(p+1))/(xgridm(k)-1);
end

```



```

end

for i=p+2:length(xgridm)
    nn(i,i)=(xgridm(i)-xgridm(i-1))/(xgridm(i)-1));
    for k=i+1:length(xgridm)
        nn(i,k)=(xgridm(i+1)-xgridm(i))/(xgridm(k)-1)+(xgridm(i)-
xgridm(i-1))/(xgridm(k)-1));
    end
end

%% Run the Simulation
tstart=0; global tfinal; tfinal=600;
tgrid1(:,1)=tstart:tfinal; %simulation time

options=odeset('NonNegative',1:Nb,'Stats','on',...
    'Refine',8,'AbsTol',1e-16,'RelTol',1e-6);

[t,y] = ode15s(@(t,y) combined_fp_ode(t,y,S_g,S_a,Nb,n,nn,p),
tgrid1,...
    mole,options);

```

```

function dCdt=combined_fp_ode(t,y,S_g,S_a,Ng,n,nn,p)

%State variables
C=zeros(Ng,1);
for i=1:Ng
    C(i)=y(i);
end

%ODEs
dC=zeros(1,Ng);

summation=0;
for j=2:Ng
    summation=summation+S_a(j)*nn(1,j)*C(j);
end

dC(1)=2*S_g(2)*C(2)+sum(S_g(3:Ng).*C(3:Ng))+summation;

for i=2:p
    summation=0;
    for j=i+1:Ng
        summation=summation+S_a(j)*nn(i,j)*C(j);
    end
    dC(i)=S_g(i+1)*C(i+1)-S_g(i)*C(i)+summation-S_a(i)*C(i);
end

summation=0;
for j=p+2:Ng
    summation=summation+S_a(j)*nn(p+1,j)*C(j);
end
dC(p+1)=n(p+1,p+2)*S_g(p+2)*C(p+2)-S_g(p+1)*C(p+1)+summation...
-S_a(p+1)*C(p+1);

for i=p+2:Ng-1
    summation=0;
    for j=i:Ng
        summation=summation+S_a(j)*nn(i,j)*C(j);
    end
    dC(i)=n(i,i)*S_g(i)*C(i)+n(i,i+1)*S_g(i+1)*C(i+1)-S_g(i)*C(i)...
+summation-S_a(i)*C(i);
end

dC(Ng)=n(Ng,Ng)*S_g(Ng)*C(Ng)-S_g(Ng)*C(Ng)...
+S_a(Ng)*nn(Ng,Ng)*C(Ng)-S_a(Ng)*C(Ng);

%Output the state variables
dCdt=dC';

```

C.2 MATLAB Code for The Simulation of Simultaneous Random and Chain-End Scissions Using The Exact Solution

```

%% Meshing
Mn_avg=4100; %number-average DP
Mw_avg=5430; %weight-average DP
xs=1; %monomer DP
xl=22496; %largest polymer DP
omega=Mn_avg/(Mw_avg-Mn_avg);
kappa=Mw_avg-Mn_avg;
Mtot=10; %initial starch concentration (g/L)

xgridl=xs:1:xl; xgridl=xgridl(:); %arithmetetic meshing

Nbl=length(xgridl); %number of pivot points
x_avgl=zeros(Nbl-1,1); i=1:Nbl-1; x_avgl(i,1)=(xgridl(i)+xgridl(i+1))/2;
x_avgl=[xgridl(1);x_avgl;xgridl(end)]; %v(i)

%% Initial MWD

%intermediate to calculate mass_tot
PP=@(x) (((x-xs)./beta).^(alpha-1).*exp(-(x-xs)./beta))./...
    (beta.*gamma(alpha)).*(x.*162+18)-(xs.*162+18));

% molar concentration density (mol/L/DP)
mass_tot=quad(PP,x_avgl(1),x_avgl(end));
P=@(x) (((x-xs)./beta).^(alpha-1).*exp(-(x-xs)./beta))./...
    (beta.*gamma(alpha)).*Mtot./mass_tot;

% molar concentration (mol/L)
mole=zeros(Nbl,1);
for i=1:Nbl
    mole(i,1)=quad(P,x_avgl(i),x_avgl(i+1));
end

%% Rate Kernel for Chain-End Scission
kp_g=1/80;
S_g1=kp_g*(xgridl).^1;

%% Rate Kernel for Random Scission
kp_a=1/200;
S_al=kp_a*(xgridl-1).^1;

%% Chunk by chunk solving
tstart=0; global tfinal; tfinal=600;
tgrid(:,1)=tstart:tfinal;

yn=mole;
yy=[];
ttime=[];
stats_exact=[];

msz=1184; %S - size of each partition

```

```

mrt{1}=zeros(msz,msz);
for j=1:Nbl/msz
    global ode_count;
    ode_count=j;
    i1=Nbl-msz*j+1;
    i2=Nbl-msz*(j-1);

    if j<Nbl/msz
        n_full=zeros(msz,j*msz);
        for i=1:msz
            for jj=i+1:j*msz
                n_full(i,jj)=(2/(Nbl-msz*j+jj-1))*S_al(Nbl-msz*j+jj);
            end
            n_full(i,i)=n_full(i,i)-S_al(Nbl-msz*j+i)-S_gl(Nbl-
msz*j+i);
        end
        for i=1:msz-1
            n_full(i,i+1)=n_full(i,i+1)+S_gl(Nbl-msz*j+i+1);
        end
        if j~=1
            n_full(msz,msz+1)=n_full(msz,msz+1)+S_gl(Nbl-
msz*j+msz+1);
        end
    elseif j==Nbl/msz
        n_full=zeros(msz,j*msz);
        for i=1:msz
            for jj=i+1:j*msz
                n_full(i,jj)=(2/(Nbl-msz*j+jj-1))*S_al(Nbl-msz*j+jj);
            end
        end
        n_full(1,2)=n_full(1,2)+2*S_gl(2);
        for jj=3:j*msz
            n_full(1,jj)=n_full(1,jj)+S_gl(jj);
        end
        for i=2:msz-1
            n_full(i,i+1)=n_full(i,i+1)+S_gl(Nbl-msz*j+i+1);
            n_full(i,i)=n_full(i,i)-S_al(Nbl-msz*j+i)-S_gl(Nbl-
msz*j+i);
        end
        n_full(msz,msz)=n_full(msz,msz)-S_al(Nbl-msz*j+msz)...
            -S_gl(Nbl-msz*j+msz);
        n_full(msz,msz+1)=n_full(msz,msz+1)+S_gl(Nbl-msz*j+msz+1);
    end

    nrt{j}=n_full(:,1:msz);
    if j>1
        mrt{j}=n_full(:,msz+1:end);
    end

    n_full=[];

    options=odeset('NonNegative',1:length(yn(i1:i2)),'Stats','on
'Refine',8,'AbsTol',1e-16,'RelTol',1e-6,...
'JPattern',triu(ones(size(nrt{j},1),size(nrt{j},2))));

    sol = ode15s(@(t,y) combined_exact_ode(t,y,nrt{j},mrt{j}),
ttime,...
yy,j,msz),tgrid,yn(i1:i2),options);

    nrt{j}=[];
    mrt{j}=[];

    stats_exact{j}=sol.stats;
    ttime{j} = sol.x';
    yy{j}=sol.y';
end

```

```

function ypr = combined_exact_ode(t,y,nrt,mrt,ttime,yy,j,msz)

yy1 = zeros(msz,1);

if j>1
    yy1=[];
    for ii=1:j-1
        yysol=interp1(ttime{ii},yy{ii},t,'linear');
        yy1=[yysol yy1];
    end
end

ypr = nrt*y(:)+mrt*yy1(:);

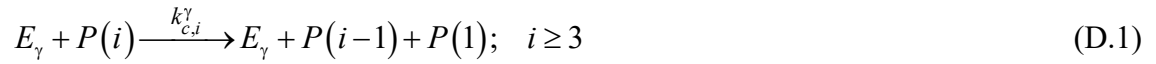
ypr=ypr(:);

```

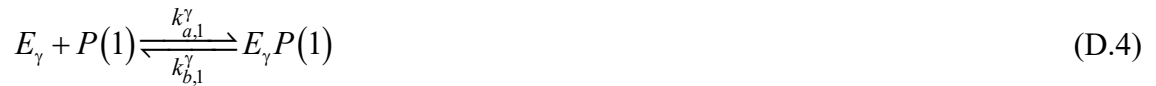
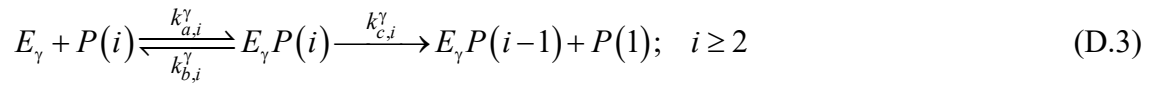
APPENDIX D: EXAMPLES OF PBM IMPLEMENTATION ON ENZYMATIC HYDROLYSIS

The specific form of PBEs in Section 5.1.1 to be solved is dependent on the mechanisms of enzyme-substrate interaction assumed. To illustrate, consider a single enzyme system with the following enzyme-substrate interactions that leads to chain-end scission:

Case A:



Case B:



where E_γ represents an enzyme that exhibits chain-end scission, $P(i)$ is a polymer population with $\text{DP} = i$, $E_\gamma P(i)$ is the enzyme bounded $P(i)$, and $k_{a,i}^\gamma$, $k_{b,i}^\gamma$, $k_{c,i}^\gamma$ (i.e. elements of the vector \mathbf{k}) are the rate constants for $\text{DP} = i$. In Case A, an irreversible scission mechanism is assumed by ignoring the formation of enzyme complexes and the cleaved components are released into the broth resulting in the recovery of the free enzyme. For Case B, bond scission occurs following the reversible formation of an enzyme-substrate complex and the enzyme exhibits a multiple attack pattern (Kurasin & Valjamae, 2011) where the bounded polymer is not released until fully cleaved into the monomer. Both mechanisms are reflected in the following different sets of PBEs, with

C_i and $C_{B,i}^\gamma$ being the molar concentration of the free and the bounded species respectively:

Case A:

$$\frac{dC_1}{dt} = \underbrace{2k_{c,2}^\gamma E_\gamma C_2 + \sum_{j=3}^{p+q} k_{c,j}^\gamma E_\gamma C_j}_{\text{Enzymatic Generation}} \quad (\text{D.5})$$

$$\frac{dC_i}{dt} = \underbrace{\sum_{j=i}^{i+1} n_{ij}^\gamma k_{c,j}^\gamma E_\gamma C_j}_{\text{Enzymatic Generation}} - \underbrace{k_{c,i}^\gamma E_\gamma C_i}_{\text{Enzymatic Scission}} \quad ; \quad i = 2 \text{ to } p+q \quad (\text{D.6})$$

$$E_\gamma = E_\gamma(0) \quad (\text{D.7})$$

Case B:

$$\frac{dC_1}{dt} = \underbrace{k_{b,1}^\gamma C_{B,1}^\gamma + \sum_{j=2}^{p+q} k_{c,j}^\gamma C_{B,j}^\gamma}_{\text{Enzymatic Generation}} - \underbrace{k_{a,1}^\gamma E_\gamma C_1}_{\text{Enzymatic Scission}} \quad (\text{D.8})$$

$$\frac{dC_{B,1}^\gamma}{dt} = \underbrace{k_{a,1}^\gamma E_\gamma C_1 + k_{c,2}^\gamma C_{B,2}^\gamma}_{\text{Enzymatic Generation}} - \underbrace{k_{b,1}^\gamma C_{B,1}^\gamma}_{\text{Enzymatic Scission}} \quad (\text{D.9})$$

$$\frac{dC_i}{dt} = \underbrace{k_{b,i}^\gamma C_{B,i}^\gamma}_{\text{Enzymatic Generation}} - \underbrace{k_{a,i}^\gamma E_\gamma C_i}_{\text{Enzymatic Scission}} \quad ; \quad i = 2 \text{ to } p+q \quad (\text{D.10})$$

$$\frac{dC_{B,i}^\gamma}{dt} = \underbrace{k_{a,i}^\gamma E_\gamma C_i + \sum_{j=i}^{i+1} n_{ij}^\gamma k_{c,j}^\gamma C_{B,j}^\gamma}_{\text{Enzymatic Generation}} - \underbrace{\left[k_{b,i}^\gamma C_{B,i}^\gamma + k_{c,i}^\gamma C_{B,i}^\gamma \right]}_{\text{Enzymatic Scission}} \quad ; \quad i = 2 \text{ to } p+q \quad (\text{D.11})$$

$$E_\gamma = E_\gamma(0) - \sum_{j=1}^{p+q} C_{B,j}^\gamma \quad (\text{D.12})$$

In the above, $n_{ii}^\gamma = 0$ for $i = 2$ to $p+1$, $n_{i,i+1}^\gamma = 1$ for $i = 2$ to p , and $n_{p+q,p+q+1}^\gamma = 0$ according to the results established in Chapter 4. The remaining n_{ij}^γ 's are given by:

$$n_{ij}^\gamma = \int_{x_i}^{x_{i+1}} \left[\frac{x_{i+1} - v}{x_{i+1} - x_i} \right] \delta(v - [x_j - v_m]) dv + \int_{x_{i-1}}^{x_i} \left[\frac{v - x_{i-1}}{x_i - x_{i-1}} \right] \delta(v - [x_j - v_m]) dv \quad (\text{D.13})$$

APPENDIX E: KINETICS OF MICROBIAL FERMENTATION

E.1 Examples of Kompala's Cybernetic Model Written in The General Form

The form of CM framework given in Section 5.1.2 is a general one represented by vector and matrix notations. The more recent variants of the cybernetic models were all given in this form (D. Ramkrishna & Song, 2012; Song et al., 2009; Song & Ramkrishna, 2011) due to the relatively more elaborate structure. The earlier versions of cybernetic models, however, can also be recast into this general form. For illustration, the cybernetic model developed by Kompala and Ramkrishna (1986), in the general form, is as follow:

Biomass growth:

$$\frac{1}{X} \frac{dX}{dt} = \mu; \quad \mu = \|\mathbf{r}\|_1; \quad \mathbf{r} = \underbrace{\begin{bmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & V_{n_V} \end{bmatrix}}_{\mathbf{D}(\mathbf{V})} \underbrace{\begin{bmatrix} r_{X,1} \\ r_{X,2} \\ \vdots \\ r_{X,n_V} \end{bmatrix}}_{\mathbf{r}_X}; \quad r_{X,i} = \frac{\mu_{\max,i} \left(\frac{e_i}{e_{\max,i}} \right) \mathcal{N}_i}{K_i + \mathcal{N}_i} \quad (\text{E.1})$$

Depletion of nutrients:

$$\frac{1}{X} \frac{d\mathcal{N}}{dt} = \underbrace{\mathbf{W}_{\mathcal{N}}}_{\mathbf{W}_{\mathcal{N}}} \mathbf{r} = \underbrace{\begin{bmatrix} -1/Y_1 & 0 & \cdots & 0 \\ 0 & -1/Y_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1/Y_{n_V} \end{bmatrix}}_{\mathbf{W}_{\mathcal{N}}} \underbrace{\begin{bmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & V_{n_V} \end{bmatrix}}_{\mathbf{r}=\mathbf{D}(\mathbf{V})\mathbf{r}_X} \underbrace{\begin{bmatrix} r_{X,1} \\ r_{X,2} \\ \vdots \\ r_{X,n_V} \end{bmatrix}}_{\mathbf{r}_X} \quad (\text{E.2})$$

$$\mathcal{N} = [\mathcal{N}_1 \quad \mathcal{N}_2 \quad \cdots \quad \mathcal{N}_{n_V}]^T$$

Synthesis of key metabolic enzymes:

$$\begin{aligned}
 \frac{de}{dt} &= \rho_e + D(U)r_e - [D(\beta) + \mu I]e & r_{e,i} &= \frac{k_{e,i}\mathcal{N}_i}{K_{e,i} + \mathcal{N}_i} \\
 &= \mathbf{0} + \underbrace{\begin{bmatrix} U_1 & 0 & \cdots & 0 \\ 0 & U_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_{n_{\mathcal{N}}} \end{bmatrix}}_{D(U)} \underbrace{\begin{bmatrix} r_{e,1} \\ r_{e,2} \\ \vdots \\ r_{e,n_{\mathcal{N}}} \end{bmatrix}}_{r_e} \\
 &\quad - \left\{ \underbrace{\begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \beta_{n_{\mathcal{N}}} \end{bmatrix}}_{D(\beta)} + \mu \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \right\} \underbrace{\begin{bmatrix} e_1/e_{\max,1} \\ e_2/e_{\max,2} \\ \vdots \\ e_{n_{\mathcal{N}}}/e_{\max,n_{\mathcal{N}}} \end{bmatrix}}_e
 \end{aligned} \tag{E.3}$$

Cybernetic variables:

$$\begin{aligned}
 U_i &= \frac{r_{X,i}X}{\sum_j r_{X,j}X} = \frac{r_{X,i}}{\sum_j r_{X,j}} \\
 V_i &= \frac{r_{X,i}X}{\max(r_{X,1}X, \dots, r_{X,j}X)} = \frac{r_{X,i}}{\max(r_{X,1}, \dots, r_{X,j})}; \quad i = 1 \text{ to } n_{\mathcal{N}}
 \end{aligned} \tag{E.4}$$

Here, $n_{\mathcal{N}}$ is the number of consumable substrates. Their simple model focused mainly on the prediction of cell growth and substrate depletion profiles, therefore the concentrations of intracellular and extracellular metabolites were not accounted for. In addition, the constitutive rate of enzyme synthesis vanished in Eq. (E.3), i.e. $\rho_{e,i} = 0$.

E.2 Rationale for Not Equating $k_{e,i} = \mu_{\max,i} + \beta_{e,i}$

In the paper by Kompala and Ramkrishna (1986), the rate constant for the synthesis of enzyme was given as $k_{e,i} = \mu_{\max,i} + \beta_{e,i}$. This form of expression had been used subsequently by other authors (Altintas et al., 2002; Kroumov et al., 2006; Ochoa et al., 2007). This appendix attempts to show that this form of expression should only be used under specific assumptions. Consider the enzyme equation of the form displayed by Eq. (5.29) in Section 5.1.3:

$$\frac{d}{dt} \left[\frac{e_i(t)}{e_{\max,i}} \right] = r_{e,i} U_{e,i} - [\mu(t) + \beta_{e,i}] \left[\frac{e_i(t)}{e_{\max,i}} \right]; \quad r_{e,i} = \frac{k_{e,i} M_i C_i(t)}{K_{e,i} + M_i C_i(t)} \quad (\text{E.5})$$

where for simplicity, only free substrates are considered here. When $e_i = e_{\max,i}$, Eq. (E.5) becomes:

$$0 = \frac{k_{e,i} M_i C_i(t)}{K_{e,i} + M_i C_i(t)} - [\mu(t) + \beta_{e,i}] \quad (\text{E.6})$$

The cybernetic variable for the induction of enzyme synthesis, $U_{e,i} = 1$ in this case because otherwise $e_{\max,i}$ would not be achieved. Following this, Eq. (E.6) can be rearranged as follows:

$$\frac{k_{e,i} M_i C_i(t)}{K_{e,i} + M_i C_i(t)} - \mu(t) = \beta_{e,i} \quad (\text{E.7})$$

As $\mu(t) = [1/X(t)] [dX(t)/dt]$, and using the relationship $dX(t)/dt = X(t) \sum_j r_{X,j} V_j$, it

follows that:

$$\frac{k_{e,i} M_i C_i(t)}{K_{e,i} + M_i C_i(t)} - \frac{\mu_{\max,i} M_i C_i(t)}{K_i + M_i C_i(t)} = \beta_{e,i}; \quad \because r_{X,i} = \frac{\mu_{\max,i} M_i C_i(t)}{K_i + M_i C_i(t)} \text{ and } V_i = 1 \text{ and } V_{j \neq i} = 0 \quad (\text{E.8})$$

Here, $V_i = 1$ and $V_{j \neq i} = 0$ following the rationale given for $U_{e,i} = 1$ above. Referring to

Eq. (E.8), if $K_{e,i} = K_i$, then:

$$\frac{M_i C_i(t)}{K_i + M_i C_i(t)} [k_{e,i} - \mu_{\max,i}] = \beta_{e,i} \quad (\text{E.9})$$

From Eq. (E.9), if the value of $K_i \ll M_i C_i(t)$, then $k_{e,i} = \mu_{\max,i} + \beta_{e,i}$. The assumption of $K_{e,i} = K_i$ was used in the work of Kompala and Ramkrishna (1986) for simplicity. However, unless confirmed experimentally, it is probable that their values may differ. In this case, it would be more appropriate to retain $k_{e,i}$ as it is in Eq. (E.5) and to avoid correlating it to $\mu_{\max,i}$ and $\beta_{e,i}$.

APPENDIX F: EXTRACELLULAR DEPOLYMERASE PRODUCTION

F.1 Derivation of The Extracellular Enzyme Production Equation

Given that m_E is the mass concentration of an extracellular enzyme within the broth and X is the mass concentration of the biomass in the reactor, the material balance for the mass of extracellular enzyme within the broth can be written as:

$$\frac{d}{dt}[V_R m_E(t)] = V_R X(t) \left[\rho_e + r_e U_e - \beta_e \frac{m_E(t)}{X(t)} \right] \quad (\text{F.1})$$

where V_R is the volume of the reactor, ρ_e is the rate of constitutive enzyme synthesis, r_e is the inducible rate of enzyme synthesis, U_e is the cybernetic variable for inducing the production of the enzyme, and β_e is the deactivation rate constant. Assuming that V_R is a constant, Eq. (F.1) becomes:

$$\frac{dm_E(t)}{dt} = \rho_e X(t) + r_e U_e X(t) - \beta_e m_E(t) \quad (\text{F.2})$$

This is the equation for describing the rate of change in the mass concentration of an extracellular enzyme within the broth. Given that M_E is the molecular weight of the enzyme, Eq. (F.2) can be re-written in terms of E^0 , i.e. the molar concentration of the enzyme within the broth:

$$\frac{dE^0(t)}{dt} = \rho_E X(t) + r_E U_E X(t) - \beta_E E^0(t); \quad \rho_E = \frac{\rho_e}{M_E}; \quad r_E = \frac{r_e}{M_E} \quad (\text{F.3})$$

The form of Eq. (F.3) is essentially that displayed by Eq. (5.15) in Section 5.1.3 where the subscript ‘ e ’ is changed to ‘ E ’ to reflect its extracellular representation. Since extracellular enzymes are excreted from within the cells, it should be fairly easy to correspondingly derive the concentration of the enzyme within the cells prior to its

excretion. Ignoring the effect of mass transfer, the concentration of an enzyme within the cells can be written as:

$$\begin{aligned} \frac{d}{dt} \left[\frac{m_E(t)}{X(t)} \right] &= m_E(t) \frac{d}{dt} \left[\frac{1}{X(t)} \right] + \frac{1}{X(t)} \frac{dm_E(t)}{dt} \\ &= -\frac{m_E(t)}{X(t)} \frac{1}{X(t)} \frac{dX(t)}{dt} + \frac{1}{X(t)} \frac{dm_E(t)}{dt} \end{aligned} \quad (\text{F.4})$$

Substituting Eq. (F.2) into the second term on the RHS of Eq. (F.4) results in the following expression:

$$\frac{de(t)}{dt} = \rho_e + r_e U_e - (\beta_e + \mu) e(t); \quad \mu = \frac{1}{X(t)} \frac{dX(t)}{dt}; \quad e(t) = \frac{m_E(t)}{X(t)} \quad (\text{F.5})$$

This is essentially Eq. (5.14) in Section 5.1.3.

F.2 Formulation for r_j ($j = \gamma$ or α)

In Sections 5.2.1 and 5.3.1, yeast was assumed to “grow” indirectly on starch by excreting either glucoamylase (γ) or both α -amylase (α) and glucoamylase (γ), having a growth rate ($r_j X, j = \gamma$ or α). A plausible expression for r_j will be proposed here. Assuming that the i -mer can only form one type of enzyme complex, thereby resulting in the expression of only C_i and $C_{B,i}^m$ (m = all depolymerases, e.g. $m = \alpha, \gamma$ if both α -amylase and glucoamylase are present) in the broth, the traditional way of writing the rate of biomass growth $r_{X,i} X$ [g/L/h] on i with the necessary inhibition terms is as follows:

$$r_{X,i} X = \frac{\mu_{\max,i} \left[\frac{e_i}{e_{\max,i}} \right] \left\{ M_i \left[C_i + \sum_m C_{B,i}^m \right] \right\} X}{\left\{ K_i + M_i \left[C_i + \sum_m C_{B,i}^m \right] \right\} \prod_k (1 + I_k / K_{INH,k})}; \quad m = \text{all depolymerases} \quad (\text{F.6})$$

where I_k is the mass concentration of the inhibiting k species and $K_{INH,k}$ is the corresponding constant of inhibition. If the microbe could also grow on substrates $i+1$,

$i+2$, ... etc., one could sum up individual rates expressed by Eq. (F.6). A cruder approximation assumes that the kinetic parameters and relative enzyme concentrations are similar, thus allowing the substrate concentration terms to be replaced by summations. To extend Eq. (F.6) to starch, the latter is a more pragmatic approach as otherwise a large number of parameters are involved. Further, in the “growth” on starch, both ethanol and glucose are considered to be the inhibiting substances. Based on these considerations and simplifications, r_jX ($j = \gamma$ or α) could be written as:

$$r_jX = \frac{\mu_{\max,j} \left[\frac{e_j}{e_{\max,j}} \right] \left\{ \sum_{i=2}^{p+q} M_i \left[C_i + \sum_m C_{B,i}^m \right] \right\} X}{\left\{ K_j + \sum_{i=2}^{p+q} M_i \left[C_i + \sum_m C_{B,i}^m \right] \right\} \left[1 + A/K_{EtOH} \right] \left[1 + \left(M_1 \left[C_1 + \sum_m C_{B,1}^m \right] / K_I^j \right) \right]} \quad (F.7)$$

Further, according to Kompala and Ramkrishna (1986), $\mu_{\max,j} = \mu_j e_{\max,j}$. Substituting this into Eq. (F.7) yields:

$$r_jX = \frac{\mu_j e_j X \left\{ \sum_{i=2}^{p+q} M_i \left[C_i + \sum_m C_{B,i}^m \right] \right\}}{\left\{ K_j + \sum_{i=2}^{p+q} M_i \left[C_i + \sum_m C_{B,i}^m \right] \right\} \left[1 + A/K_{EtOH} \right] \left[1 + \left(M_1 \left[C_1 + \sum_m C_{B,1}^m \right] / K_I^j \right) \right]} \quad (F.8)$$

Since e_jX [g/L] is the concentration of depolymerase in the broth, Eq. (F.8) can be simplified to yield:

$$r_j = \frac{\mu_j M_j \left[E_j^0 / X \right] \left\{ \sum_{i=2}^{p+q} M_i \left[C_i + \sum_m C_{B,i}^m \right] \right\}}{\left\{ K_j + \sum_{i=2}^{p+q} M_i \left[C_i + \sum_m C_{B,i}^m \right] \right\} \left[1 + A/K_{EtOH} \right] \left[1 + \left(M_1 \left[C_1 + \sum_m C_{B,1}^m \right] / K_I^j \right) \right]}; \quad E_j^0 = \frac{e_j X}{M_j} \quad (F.9)$$

This is the form of Eq. (5.33) in Section 5.2.1 and Eqs. (5.59) - (5.60) in Section 5.3.1.

APPENDIX G: SUBSITE THEORY

G.1 Subsite Theory for Glucoamylase

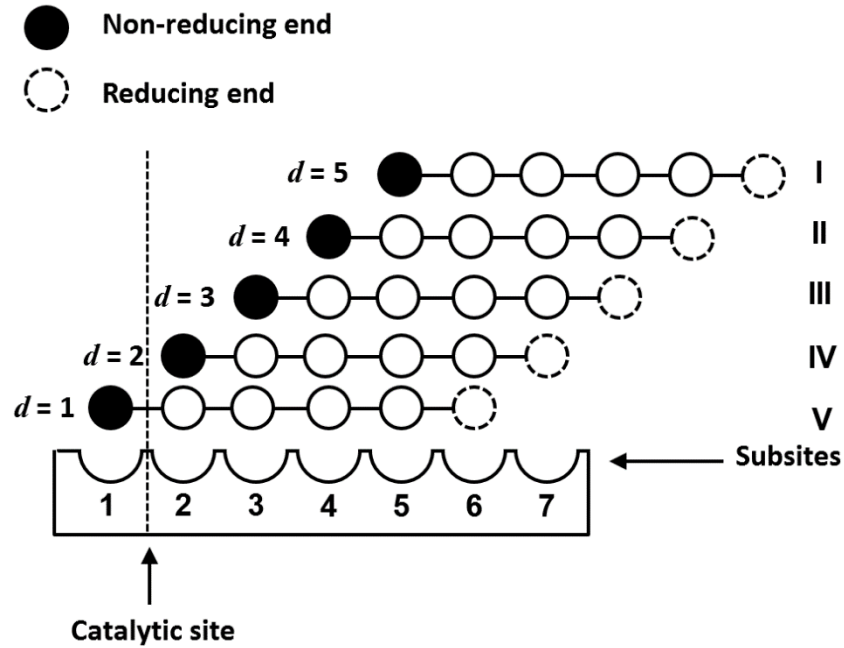


Figure G1: Different modes of binding between a DP6 oligosaccharide with the subsites of glucoamylase. The position of the non-reducing end is given by d .

The subsite theory for glucoamylase was developed by Hiromi (1970). Referring to Figure G1, the catalytic site of glucoamylase is situated between the first and the second subsite, where only binding mode V produces glucose and is referred to as the productive complex. The rest of the binding modes do not produce glucose and are referred as non-productive complexes. According to the subsite theory, the macroscopic M-M parameters, i.e. $k_{c,i}^\gamma$ and $K_{m,i}^\gamma$ in $-dC_i/dt = k_{c,i}^\gamma E_\gamma C_i / [K_{m,i}^\gamma + C_i]$, for the hydrolysis of a maltooligosaccharide by glucoamylase can be calculated using the following expressions (Hiromi, 1970; Hiromi et al., 1973):

$$\frac{1}{K_{m,i}^\gamma} = \sum_P K_{P,i}^\gamma + \sum_Q K_{Q,i}^\gamma = \sum_{J=1}^{N_m^\gamma} K_{J,i}^\gamma; \quad K_{J,i}^\gamma = \exp\left[-\frac{2 \ln(55.5)}{1000R}\right] \exp\left[\sum_n^{\text{occ.}} \frac{A_n^\gamma}{RT}\right]_{J,i} \quad (\text{G.1})$$

$$k_{c,i}^\gamma = k_{\text{int}}^\gamma \sum_P K_{P,i}^\gamma / \sum_J K_{J,i}^\gamma \quad (\text{G.2})$$

where $K_{P,i}^\gamma$ and $K_{Q,i}^\gamma$ (generally represented by $K_{J,i}^\gamma$ [mol/L]) are the association constants of the i -mer substrate in a binding mode specified by the subscript P (productive), Q (non-productive), or J (either productive or non-productive) and N_m^γ is the total number of subsites of the glucoamylase. Here, A_n^γ [kcal/mole] is the subsite affinity of the n -th subsite for glucoamylase, expressed in free energy units, and $\sum_n^{\text{occ.}}$ implies that the sum is taken for the occupied subsites. In addition, R [kcal/K/mol] and T [K] are the universal gas constant and the absolute temperature respectively whereas k_{int}^γ [1/h] is the intrinsic rate of scission for the productive complex. In other words, given the number of subsites and the subsite affinities, $K_{m,i}^\gamma$ and $k_{c,i}^\gamma$ can be calculated theoretically.

G.2 Subsite Theory for α -amylase

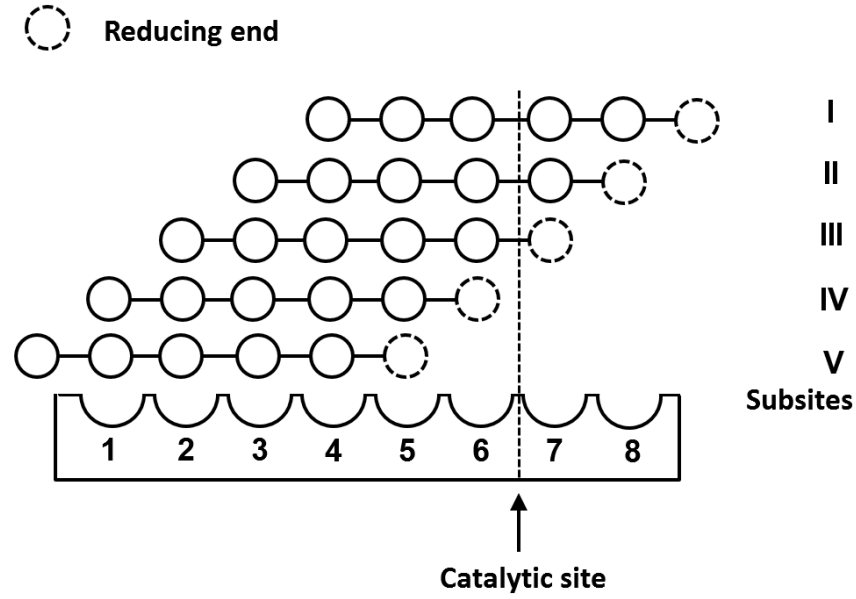


Figure G2: Different modes of binding between a DP6 oligosaccharide with the subsites of α -amylase.

The subsite theory for depolymerase was developed by Hiromi (1970) and implemented by Iwasa et al. (1974) on α -amylase. Referring to Figure G2, the catalytic site of α -amylase is situated between the sixth and the seventh subsite, where only binding modes I – III produce bond cleavage and are referred to as the productive complexes. The rest of the binding modes do not produce bond cleavage and are referred as non-productive complexes. According to the subsite theory, the macroscopic M-M parameters, i.e. $k_{c,i}^\alpha$ and $K_{m,i}^\alpha$ in $-dC_i/dt = k_{c,i}^\alpha E_\alpha C_i / [K_{m,i}^\alpha + C_i]$, for the hydrolysis of a maltooligosaccharide by α -amylase can be calculated using the following expressions (Iwasa et al., 1974):

$$\frac{1}{K_{m,i}^\alpha} = \sum_P K_{P,i}^\alpha + \sum_Q K_{Q,i}^\alpha = \sum_{J=1}^{N_m^\alpha} K_{J,i}^\alpha; \quad K_{J,i}^\alpha = \exp \left[-\frac{2 \ln(55.5)}{1000R} \right] \exp \left[\sum_n^{\text{occ.}} \frac{A_n^\alpha}{RT} \right]_{J,i} \quad (\text{G.3})$$

$$k_{c,i}^\alpha = k_{\text{int}}^\alpha \sum_P K_{P,i}^\alpha / \sum_J K_{J,i}^\alpha \quad (\text{G.4})$$

where $K_{P,i}^\alpha$ and $K_{Q,i}^\alpha$ (generally represented by $K_{J,i}^\alpha$ [mol/L]) are the association constants of the i -mer substrate in a binding mode specified by the subscript P (productive), Q (non-productive), or J (either productive or non-productive) and N_m^α is the total number of subsites of the α -amylase. Here, A_n^α [kcal/mole] is the subsite affinity of the n -th subsite for α -amylase, expressed in free energy units, and $\sum_n^{\text{occ.}}$ implies that the sum is taken for the occupied subsites. In addition, R [kcal/K/mol] and T [K] are the universal gas constant and the absolute temperature respectively whereas k_{int}^α [1/h] is the intrinsic rate of scission for the productive complex. In other words, given the number of subsites and the subsite affinities, $K_{m,i}^\alpha$ and $k_{c,i}^\alpha$ can be calculated theoretically.

G.3 Subsite Affinity Maps

Values of Subsite Affinities used in Case Study I (Section 5.2):

Table G1: Values of subsite affinities at 25 °C and pH 4.5 for glucoamylase from *R. delemar* given by Hiromi et al. (1973). The catalytic site is situated between the first and the second subsite.

Subsite	Subsite affinity (kcal/mol)
1	0.00
2	4.85
3	1.59
4	0.43
5	0.22
6	0.11
7	0.10

Values of Subsite Affinities used in Case Study II (Section 5.3):

Table G2: Values of subsite affinities for α -amylase at 25 °C and pH 5.85 from *B. subtilis* given by Iwasa et al. (1974). For the subsites adjacent to the catalytic site, i.e. subsites 6 and 7, the authors reported that the total subsite affinity = -3.1 kcal/mol. Thus, the values shown here were calculated by assuming the subsite affinity for subsite 7 = 3.1 kcal/mol (as reported by Marchal et al. (2003)).

Subsite	Subsite affinity (kcal/mol)
1	1.1
2	2.4
3	0.0
4	0.6
5	2.4
6	-6.2
7	3.1
8	1.2

Table G3: Values of subsite affinities for glucoamylase at 50 °C and pH 4.4 from *A. awamori* given by Sierks (1988). The catalytic site is situated between the first and the second subsite.

Subsite	Subsite affinity (kcal/mol)
1	−0.53
2	5.11
3	1.50
4	0.43
5	0.38
6	0.24
7	0.072

APPENDIX H: MATLAB CODE FOR THE SIMULATION OF THE SSF OF A GLUCOAMYLASE PRODUCING RECOMBINANT YEAST

The following gives the MATLAB code for Case Study I of Chapter 5:

```
%% Meshing
Mn_avg=160; %number-average DP
Mw_avg=212; %weight-average DP
xs=1; %monomer DP
xl=878; %largest polymer DP
omega=Mn_avg/(Mw_avg-Mn_avg);
kappa=Mw_avg-Mn_avg;
Mtot=30; %initial mass concentration of starch (g/L)
starchmindp=2; %minimum DP of inductible species
starchmindpl=41; %minimum DP of starch

% calculating the minimum value of p
vm=1;
total_pivots=200;
p=0; n_pivots=0;
while n_pivots < total_pivots
    p=p+1;
    n_pivots=(p+1)+log(xl/(p+vm))/log((p+2*vm)/(p+vm));
end

xgrid=1:1:p+1; xgrid=xgrid(:); %discrete region x(1) to x(p+1)

N_con=total_pivots-(p+1); %number of pivots in the continuous region
ratio=exp((1/N_con)*log((xl/xgrid(end)))); %geometric ratio

% geometric meshing
i=0; i=1:N_con+1;
xgrid_g=xgrid(end).*ratio.^(i-1);
xgrid=[xgrid(1:end-1);xgrid_g'];
xgrid(end)=xl;

Nb=length(xgrid); %number of pivot points
x_avg=zeros(Nb-1,1); i=1:Nb-1; x_avg(i,1)=(xgrid(i)+xgrid(i+1))/2;
x_avg=[xgrid(1);x_avg;xgrid(end)]; %v(i)

%% Initial distribution

%intermediate to calculate mass_tot
PP=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
    (kappa.*gamma(omega)).*(x.*162+18)-(xs.*162+18));

% molar distribution (mol/L/DP)
mass_tot=quad(PP,x_avg(1),x_avg(end));
P=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
    (kappa.*gamma(omega)).*Mtot./mass_tot);

% molar concentration (mol/L)
p_init=zeros(Nb,1); %unbounded species
pb_init=zeros(Nb,1); %bounded species
for i=1:Nb
    p_init(i,1)=quad(P,x_avg(i),x_avg(i+1));
end
```

```

end

%% n(i,k) for Glucoamylase
n=zeros(length(xgrid),length(xgrid));
n(1,1)=0; vm=1;
for k=2:length(xgrid)
    if (xgrid(k)-vm)<xgrid(2) && (xgrid(k)-vm)>xgrid(1)
        n(1,k)=(xgrid(2)-(xgrid(k)-vm))/(xgrid(2)-xgrid(1));
    end
end
if sum(n(1,:))==0
    n(1,2)=1;
end
for i=2:length(xgrid)
    if (xgrid(i)-vm)<xgrid(i) && (xgrid(i)-vm)>xgrid(i-1)
        n(i,i)=(xgrid(i)-vm-xgrid(i-1))/(xgrid(i)-xgrid(i-1));
    end
    for k=i+1:length(xgrid)
        if (xgrid(k)-vm)>xgrid(i) && (xgrid(k)-vm)<xgrid(i+1)
            n(i,k)=(xgrid(i+1)-(xgrid(k)-vm))/(xgrid(i+1)-xgrid(i));
        end
        if (xgrid(k)-vm)>xgrid(i-1) && (xgrid(k)-vm)<xgrid(i)
            n(i,k)=(xgrid(k)-vm-xgrid(i-1))/(xgrid(i)-xgrid(i-1));
        end
    end
    if sum(n(i,:))==0
        n(i,i+1)=1;
    end
end

%% Model Parameters
c_init=0.1; %initial cell concentration (g/L)
EtOH_init=0; %initial ethanol concentration (mol/L)
enz_init(1,1)=0.945654297; %initial enzyme level - DP1
enz_init(2,1)=0.1; %initial enzyme level - DP2
enz_init(3,1)=0.1; %initial enzyme level - DP3
E0_init=0.010683203; %initial total glucoamylase conc (g/L)
K(1,1)=0.199469511; %monod saturation constant (g/L) - DP1
K(2,1)=3*K(1); %monod saturation constant (g/L) - DP2
K(3,1)=3*K(2); %monod saturation constant (g/L) - DP3
K(4,1)=2.992675364; %monod saturation constant (g/L) - Glucoamylase
Ks(1,1)=0.930761681; %saturation constant (g/L) - DP1
Ks(2,1)=3*Ks(1); %saturation constant (g/L) - DP2
Ks(3,1)=3*Ks(2); %saturation constant (g/L) - DP3
Ks(4,1)=2.278169147; %saturation constant (g/L) - Glucoamylase
yld(1,1)=0.431414135; %yield coefficients (g/g) - DP1
yld(2,1)=0.95*yld(1); %yield coefficients (g/g) - DP2
yld(3,1)=0.95*yld(2); %yield coefficients (g/g) - DP3
beta_c(1,1)=0.012376538; %death rate of enz (h^-1) - DP1
beta_c(2,1)=beta_c(1); %death rate of enz (h^-1) - DP2
beta_c(3,1)=beta_c(1); %death rate of enz (h^-1) - DP3

```

```

beta_c(4,1)=0.001364716; %death rate of enz (h^-1) - DP4
mu_max(1,1)=0.512039536; %max specific growth rate (h^-1) - DP1
mu_max(2,1)=0.88*mu_max(1); %max specific growth rate (h^-1) - DP2
mu_max(3,1)=0.88*mu_max(2); %max specific growth rate (h^-1) - DP3
mu_max(4,1)=0.335815126; %max specific growth rate (h^-1) - ✓
Glucoamylase
ke(1,1)=0.092198693; %rate constant for enz synthesis (h^-1) - DP1
ke(2,1)=ke(1); %rate constant for enz synthesis (h^-1) - DP2
ke(3,1)=ke(1); %rate constant for enz synthesis (h^-1) - DP3
ke(4,1)=0.248150801; %rate constant for glucoamylase synthesis (h^-1)
ME=300000; %molecular weight of glucoamylase (g/mol)
EtOHmax=1.059300826; %ethanol inhibition constant on growth
yldE(1)=0.425650138; %yield of ethanol from DP1
yldE(2)=1.05*yldE(1); %yield of ethanol from DP2
yldE(3)=1.05*yldE(2); %yield of ethanol from DP2
MEtOH=46; %molecular weight of ethanol (g/mol)
cgmax=0.293072445; %glucose repression on starch growth indicator (g)
beta_E=1.93e-4; %enz degradation constant from ethanol
eta=2; %order of ethanol inhibition on glucoamylase

a0=724.0426577; %parameter for k2
b0=954.591738; %parameter for k2
c0=196.3041053; %parameter for k2

kl_hat=5e6; %parameter for k1

%% Rate constants
load KM_value2_22496; %load the KM from subsite theory
kl_dat=zeros(xl,1);
kminusl_dat=zeros(xl,1);
k0_dat=zeros(xl,1);
for i=1:xl
    k0_dat(i,1)=(a0*i/(b0+i))+c0;
    kl_dat(i,1)=kl_hat*i;
    kminusl_dat(i,1)=kl_dat(i,1)*KM_dat(i,1);
end

k0_func=spline(1:1:xl,k0_dat);
kl_func=spline(1:1:xl,kl_dat);
kminusl_func=spline(1:1:xl,kminusl_dat);

kl=zeros(length(xgrid),1);
kminusl=zeros(length(xgrid),1);
k2=zeros(length(xgrid),1);
for ii=1:length(xgrid)
    kl(ii,1)=ppval(kl_func,xgrid(ii));
    kminusl(ii,1)=ppval(kminusl_func,xgrid(ii));
    k2(ii,1)=ppval(k0_func,xgrid(ii));
end

KM=kminusl./kl; %KM value

```



```

K2=k2./kminus1; %k2/kminus1 << 1

%%
k(1)=K(1);
k(2)=K(2);
k(3)=K(3);
k(4)=K(4);
k(5)=Ks(1);
k(6)=Ks(2);
k(7)=Ks(3);
k(8)=Ks(4);
k(9)=yld(1);
k(10)=yld(2);
k(11)=yld(3);
k(12)=mu_max(1);
k(13)=mu_max(2);
k(14)=mu_max(3);
k(15)=mu_max(4);
k(16)=a0;
k(17)=b0;
k(18)=c0;
k(19)=ke(1);
k(20)=ke(2);
k(21)=ke(3);
k(22)=ke(4);
k(23)=EtOHmax;
k(24)=yldE(1);
k(25)=yldE(2);
k(26)=yldE(3);
k(27)=cgmax;
k(28)=eta;
k(29)=beta_c(1);
k(30)=beta_c(2);
k(31)=beta_c(3);
k(32)=beta_c(4);
k(33)=beta_E;

%% Run the simulation
tstart=0; global tfinal; tfinal=200;
tgrid(:,1)=tstart:tfinal; %simulation time

options=odeset('NonNegative',...
    1:length([p_init;pb_init;enz_init;E0_init;EtOH_init;c_init;]),...
    'Stats','on','OutputFcn','odeprogress','Refine',...
    8,'AbsTol',1e-22,'RelTol',1e-6);
[t,y] = ode15s(@(t,y) pbm_cyber_ode(t,y,k1,kminus1,k2,...
    ME,MEtOH,Nb,n,k,xgrid,starchmindp),...
    tgrid,[p_init;pb_init;enz_init;E0_init;EtOH_init;c_init;],...
    options);

```

```

function dydt=pbm_cyber_ode(t,y,k1,kminus1,k2,ME,MEtOH,...
    Nb,n,k,xgrid,starchmindp)

K(1)=k(1);
K(2)=k(2);
K(3)=k(3);
K(4)=k(4);
Ks(1)=k(5);
Ks(2)=k(6);
Ks(3)=k(7);
Ks(4)=k(8);
yld(1)=k(9);
yld(2)=k(10);
yld(3)=k(11);
mu(1)=k(12);
mu(2)=k(13);
mu(3)=k(14);
mu(4)=k(15);
ke(1)=k(19);
ke(2)=k(20);
ke(3)=k(21);
ke(4)=k(22);
EtOHmax=k(23);
yldE(1)=k(24);
yldE(2)=k(25);
yldE(3)=k(26);
cgmax=k(27);
eta=k(28);
beta(1)=k(29);
beta(2)=k(30);
beta(3)=k(31);
beta(4)=k(32);
beta_E=k(33);

%State variables
g=zeros(Nb,1); %substrates (mol/L)
gb=zeros(Nb,1); %enz complexes (mol/L)
for i=1:Nb
    g(i,1)=y(i); %soluble substrate - DP(i)
    gb(i,1)=y(Nb+i); %bounded substrate - DP(i)
end
sugar(1,1)=(g(1)+gb(1))*(162*xgrid(1)+18); %total DP1 (g/L)
sugar(2,1)=(g(2)+gb(2))*(162*xgrid(2)+18); %total DP2 (g/L)
sugar(3,1)=(g(3)+gb(3))*(162*xgrid(3)+18); %total DP3 (g/L)

starch=0; %total starch (g/L)
for i=starchmindp:Nb
    starch=starch+(162*xgrid(i)+18)*g(i)+(162*xgrid(i)+18)*gb(i);
end

enz(1,1)=y(2*Nb+1); %dimensionless enzyme level - DP1

```

```

enz(2,1)=y(2*Nb+2); %dimensionless enzyme level - DP2
enz(3,1)=y(2*Nb+3); %dimensionless enzyme level - DP3
E0=y(2*Nb+4); %total extracellular glucoamylase (g/L)

EtOH=y(2*Nb+5); %ethanol (mol/L)
Ethanol_0=EtOH*MEtOH; %ethanol (g/L)

c=y(end); %biomass conc (g/L)

re=zeros(4,1); %enz synthesis rate (h^-1)
r=zeros(4,1); %biomass growth rate (g L^-1 h^-1)
for i=1:3
    re(i,1)=ke(i)*sugar(i)/(Ks(i)+sugar(i));
    r(i,1)=mu(i)*enz(i)*sugar(i)*c/((K(i)+sugar(i))*...
        *(1+(Ethanol_0/EtOHmax)));
end

re(4,1)=ke(4)*starch*c/(Ks(4)+starch); %glucoamylase induction
r(4,1)=mu(4)*E0*starch/((K(4)+starch)*(1+(Ethanol_0/EtOHmax))*...
    *(1+(sugar(1)/cgmax))); %growth indicator for glucoamylase ✓
production

%Cybernetic variables
u=r./sum(r); %dimensionless
v=r(1:3)./max(r(1:3)); %dimensionless

%Enzyme balance
E=E0/ME-sum(gb);

%ODEs
dg=zeros(1,Nb);
dgb=zeros(1,Nb);
denz=zeros(1,3);

%biomass (g L^-1 h^-1)
dc=sum(r(1:3).*v(1:3));

%ethanol (mol L^-1 h^-1)
dEtOH=sum((yldE(1:3)'./yld(1:3)').*r(1:3).*v(1:3))/MEtOH;

%DP1
dg(1)=-k1(1)*E*g(1)...
    +kminus1(1)*gb(1)+k2(2)*gb(2)+sum(k2(2:end).*gb(2:end))...
    -(r(1)*v(1)/yld(1))/(180); %soluble DP1 (mol L^-1 h^-1)
dgb(1)=k1(1)*E*g(1)-kminus1(1)*gb(1); %bounded DP1 (mol L^-1 h^-1)
denz(1)=re(1)*u(1)-beta(1)*enz(1)-(enz(1)/c)*dc; %enzyme level (h^-1)

%DP2 - DP3
for i=2:3
    summation=0;
    for j=i:Nb

```

```

        summation=summation+n(i,j)*k2(j)*gb(j);
    end
    dg(i)=-k1(i)*E*g(i)+kminusl(i)*gb(i)...
        +summation-(r(i)*v(i)/yld(i))/(162*xgrid(i)+18);
    dgb(i)=k1(i)*E*g(i)-kminusl(i)*gb(i)-k2(i)*gb(i);
    denz(i)=re(i)*u(i)-beta(i)*enz(i)-(enz(i)/c)*dc;
end

%change in extracellular glucoamylase (g L-1 h-1)
dE0=re(4)*u(4)-beta(4)*(1+beta_E*Ethanol_0^eta)*E0;

%DP4 - DPN
for i=3+1:Nb
    summation=0;
    for j=i:Nb
        summation=summation+n(i,j)*k2(j)*gb(j);
    end
    dg(i)=-k1(i)*E*g(i)+kminusl(i)*gb(i)+summation;
    dgb(i)=k1(i)*E*g(i)-kminusl(i)*gb(i)-k2(i)*gb(i);
end

%Output the state variables
dydt=[dg';dgb';denz'; dE0; dEtOH; dc;];

```

APPENDIX I: MATLAB CODE FOR THE SIMULATION OF THE SSF OF AN A-AMYLASE AND GLUCOAMYLASE PRODUCING RECOMBINANT YEAST

The following gives the MATLAB code for Case Study II of Chapter 5:

```
%% Meshing
Mn_avg=160; %number-average DP
Mw_avg=212; %weight-average DP
xs=1; %monomer DP
xl=878; %largest polymer DP
omega=Mn_avg/(Mw_avg-Mn_avg);
kappa=Mw_avg-Mn_avg;
Mtot=38.4; %initial mass concentration of starch (g/L)

% calculating the minimum value of p
vm=1;
total_pivots=200;
p=0; n_pivots=0;
while n_pivots < total_pivots
    p=p+1;
    n_pivots=(p+1)+log(xl/(p+vm))/log((p+2*vm)/(p+vm));
end

xgrid=1:l:p+1; xgrid=xgrid(:); %discrete region x(1) to x(p+1)

N_con=total_pivots-(p+1); %number of pivots in the continuous region
ratio=exp((1/N_con)*log((xl/xgrid(end)))); %geometric ratio

% geometric meshing
i=0; i=1:N_con+1;
xgrid_g=xgrid(end).*ratio.^(i-1);
xgrid=[xgrid(1:end-1);xgrid_g'];
xgrid(end)=xl;

starchmindp=2; %minimum DP of inductible species
starchdpmin=41; %minimum starch DP

Nb=length(xgrid); %number of pivot points
x_avg=zeros(Nb-1,1); i=1:Nb-1; x_avg(i,1)=(xgrid(i)+xgrid(i+1))/2;
x_avg=[xgrid(1);x_avg;xgrid(end)]; %v(i)

%% Initial distribution

%intermediate to calculate mass_tot
PP=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
    (kappa.*gamma(omega)).*((x.*162+18)-(xs.*162+18));

%molar distribution (mol/L/DP)
mass_tot=quad(PP,x_avg(1),x_avg(end));
P=@(x) (((x-xs)./kappa).^(omega-1).*exp(-(x-xs)./kappa))./...
    (kappa.*gamma(omega)).*Mtot./mass_tot;

%molar concentration (mol/L)
p_init=zeros(Nb,1); %unbounded species
pby_init=zeros(Nb,1); %bounded species by glucoamylase
pba_init=zeros(Nb,1); %bounded species by alpha-amylase
```

```

for i=1:Nb
    p_init(i,1)=quad(P,x_avg(i),x_avg(i+1));
end
p_init(1)=0.79/180;
p_init(2)=1.69/342;

%% n(i,k) for Glucoamylase
n=zeros(length(xgrid),length(xgrid));
n(1,1)=0; vm=1;
for k=2:length(xgrid)
    if (xgrid(k)-vm)<xgrid(2) && (xgrid(k)-vm)>xgrid(1)
        n(1,k)=(xgrid(2)-(xgrid(k)-vm))/(xgrid(2)-xgrid(1));
    end
end
if sum(n(1,:))==0
    n(1,2)=1;
end
for i=2:length(xgrid)
    if (xgrid(i)-vm)<xgrid(i) && (xgrid(i)-vm)>xgrid(i-1)
        n(i,i)=(xgrid(i)-vm-xgrid(i-1))/(xgrid(i)-xgrid(i-1));
    end
    for k=i+1:length(xgrid)
        if (xgrid(k)-vm)>xgrid(i) && (xgrid(k)-vm)<xgrid(i+1)
            n(i,k)=(xgrid(i+1)-(xgrid(k)-vm))/(xgrid(i+1)-xgrid(i));
        end
        if (xgrid(k)-vm)>xgrid(i-1) && (xgrid(k)-vm)<xgrid(i)
            n(i,k)=(xgrid(k)-vm-xgrid(i-1))/(xgrid(i)-xgrid(i-1));
        end
    end
    if sum(n(i,:))==0
        n(i,i+1)=1;
    end
end
end

%% n(i,k) for Alpha-amylase
xgridm=xgrid;
nn=zeros(length(xgridm),length(xgridm));
nn(1,1)=0;
i=1;
for k=2:length(xgridm)
    nn(1,k)=2/(xgridm(k)-1);
end
for i=2:p
    for k=i+1:length(xgridm)
        nn(i,k)=2/(xgridm(k)-1);
    end
end
end

for k=i+1:length(xgridm)
    nn(p+1,k)=(xgridm(p+2)-xgridm(p+1))/(xgridm(k)-1);
end
end

```



```

for i=p+2:length(xgridm)
    nn(i,i)=((xgridm(i)-xgridm(i-1))/(xgridm(i)-1));
    for k=i+1:length(xgridm)
        nn(i,k)=((xgridm(i+1)-xgridm(i))/(xgridm(k)-1))...
            +((xgridm(i)-xgridm(i-1))/(xgridm(k)-1));
    end
end

%% Model Parameters
c_init=0.1; %initial cell concentration (g/L)
EtOH_init=0; %initial ethanol concentration (mol/L)
enz_init(1,1)=0.9; %initial enzyme level - DP1
enz_init(2,1)=0.1; %initial enzyme level - DP2
enz_init(3,1)=0.1; %initial enzyme level - DP3
Ey0_init=0.042055034; %initial total glucoamylase concentration (g/L)
Ea0_init=0.14162; %initial total alpha-amylase concentration (g/L)
K(1,1)=1.215948427; %monod saturation constant (g/L) - DP1
K(2,1)=3*K(1); %monod saturation constant (g/L) - DP2
K(3,1)=3*K(2); %monod saturation constant (g/L) - DP3
K(4,1)=13.86039975; %monod saturation constant (g/L) - glucoamylase
K(5,1)=2.65661885; %monod saturation constant (g/L) - alpha-amylase
Ks(1,1)=3.1986936; %saturation constant (g/L) - DP1
Ks(2,1)=3*Ks(1); %saturation constant (g/L) - DP2
Ks(3,1)=3*Ks(2); %saturation constant (g/L) - DP3
Ks(4,1)=10.4337544; %saturation constant (g/L) - glucoamylase
Ks(5,1)=1.6432712; %saturation constant (g/L) - alpha-amylase
yld(1,1)=0.126652638; %yield coefficients (g/g) - DP1
yld(2,1)=0.95*yld(1); %yield coefficients (g/g) - DP2
yld(3,1)=0.95*yld(2); %yield coefficients (g/g) - DP3
beta_c(1,1)=0.01; %death rate of enz (h^-1) - DP1
beta_c(2,1)=beta_c(1); %death rate of enz (h^-1) - DP2
beta_c(3,1)=beta_c(1); %death rate of enz (h^-1) - DP3
beta_c(4,1)=0.0280110; %death rate of enz (h^-1) - glucoamylase
beta_c(5,1)=0.3845500; %death rate of enz (h^-1) - alpha-amylase
mu_max(1,1)=0.786943225; %maximum specific growth rate (h^-1) - DP1
mu_max(2,1)=0.88*mu_max(1); %maximum specific growth rate (h^-1) - ✓
DP2
mu_max(3,1)=0.88*mu_max(2); %maximum specific growth rate (h^-1) - ✓
DP3
mu_max(4,1)=0.418590363; %specific growth rate (h^-1) - glucoamylase
mu_max(5,1)=0.43317153; %specific growth rate (h^-1) - alpha-amylase
ke(1,1)=0.796943225; %rate constant for enz synthesis (h^-1) - DP1
ke(2,1)=0.702510038; %rate constant for enz synthesis (h^-1) - DP2
ke(3,1)=0.61940883344; %rate constant for enz synthesis (h^-1) - DP3
ke(4,1)=0.446601363; %rate constant for enz synthesis (h^-1)-✓
glucoamylase
ke(5,1)=0.81772153; %rate constant for enz synthesis (h^-1)-alpha-✓
amylase
MEy=83870; %molecular weight of glucoamylase (g/mol)
MEa=47000; %molecular weight of alpha-amylase (g/mol)

```

```

EtOHmaxg=7.004476428; %ethanol inhibition on growth rate
yldE(1)=0.429553214; %yield of ethanol from DP1
yldE(2)=1.05*yldE(1); %yield of ethanol from DP2
yldE(3)=1.05*yldE(2); %yield of ethanol from DP3
MEtOH=46; %molecular weight of ethanol (g/mol)
cgmax1=14.5687059; %glucose repression - glucoamylase (g)
cgmax2=4.2308480; %glucose repression - alpha-amylase (g)
beta_E1=0.021786757; %glucoamylase degradation constant by ethanol
Kd=0.011204338; %cell death rate constant (1/h)

a0y=734.2639943; %parameter for k2 (glucoamylase)
b0y=49.84604026; %parameter for k2 (glucoamylase)
c0y=18.11231316; %parameter for k2 (glucoamylase)

kl_haty=5000000; %parameter for k1 (glucoamylase)

a0a=643.812281717925; %parameter for k2 (alpha-amylase)
b0a=1.047194309150537e+004; %parameter for k2 (alpha-amylase)
c0a=6.415964447194186; %parameter for k2 (alpha-amylase)

kl_hata=5000000; %parameter for k1 (alpha-amylase)

%% Rate Constants for Glucoamylase
load KM_value_y; %load the KM from subsite theory
kly_dat=zeros(xl,1);
kminusly_dat=zeros(xl,1);
k0y_dat=zeros(xl,1);
for i=1:xl
    k0y_dat(i,1)=(a0y*i/(b0y+i))+c0y;
    kly_dat(i,1)=kl_haty*i;
    kminusly_dat(i,1)=kly_dat(i,1)*KM_dat(i,1);
end

k0y_func=spline(1:1:xl,k0y_dat);
kly_func=spline(1:1:xl,kly_dat);
kminusly_func=spline(1:1:xl,kminusly_dat);

kly=zeros(length(xgrid),1);
kminusly=zeros(length(xgrid),1);
k2y=zeros(length(xgrid),1);
for ii=1:length(xgrid)
    kly(ii,1)=ppval(kly_func,xgrid(ii));
    kminusly(ii,1)=ppval(kminusly_func,xgrid(ii));
    k2y(ii,1)=ppval(k0y_func,xgrid(ii));
end

KMy=kminusly./kly; %KM value
K2y=k2y./kminusly; %k2/kminus1 << 1

%% Rate Constants for Alpha-amylase
load KM_value_a; %load the KM from subsite theory

```



```

k1a_dat=zeros(x1,1);
kminus1a_dat=zeros(x1,1);
k0a_dat=zeros(x1,1);
for i=1:x1
    k0a_dat(i,1)=(a0a*i/(b0a+i))+c0a;
    k1a_dat(i,1)=k1_hata*i;
    kminus1a_dat(i,1)=k1a_dat(i,1)*KM_dat1(i,1);
end

k0a_func=spline(1:1:x1,k0a_dat);
k1a_func=spline(1:1:x1,k1a_dat);
kminus1a_func=spline(1:1:x1,kminus1a_dat);

k1a=zeros(length(xgrid),1);
kminus1a=zeros(length(xgrid),1);
k2a=zeros(length(xgrid),1);
for ii=1:length(xgrid)
    k1a(ii,1)=ppval(k1a_func,xgrid(ii));
    kminus1a(ii,1)=ppval(kminus1a_func,xgrid(ii));
    k2a(ii,1)=ppval(k0a_func,xgrid(ii));
end

KMa=kminus1a./k1a; %KM value
K2a=k2a./kminus1a; %k2/kminus1 << 1

%%
k(1)=K(1);
k(2)=K(4);
k(3)=K(5);
k(4)=Ks(1);
k(5)=Ks(4);
k(6)=Ks(5);
k(7)=yld(1);
k(8)=mu_max(1);
k(9)=mu_max(4);
k(10)=mu_max(5);
k(11)=a0y;
k(12)=b0y;
k(13)=c0y;
k(14)=a0a;
k(15)=b0a;
k(16)=c0a;
k(17)=ke(1);
k(18)=ke(2);
k(19)=ke(3);
k(20)=ke(4);
k(21)=ke(5);
k(22)=EtOHmaxg;
k(23)=yldE(1);
k(24)=Ey0_init;
k(25)=Ea0_init;

```

```

k(26)=cgmax1;
k(27)=cgmax2;
k(28)=beta_El;
k(29)=beta_c(1);
k(30)=beta_c(4);
k(31)=beta_c(5);
k(32)=Kd;

%% Run the simulation
tstart=0; global tfinal; tfinal=115;
tgrid(:,1)=tstart:tfinal; %simulation time

options=odeset('NonNegative',1:length([p_init;...
    pby_init;pba_init;enz_init;Ey0_init;Ea0_init;EtOH_init;
c_init;]),...
    'Stats','on','Refine',8,'AbsTol',1e-22,'RelTol',1e-6);
[t,y] = ode15s(@(t,y) combined_pbmcyber_ode_ulgen(t,y,kla,
kminusla,...
    k2a,kly,kminusly,k2y,MEy,MEa,MEtOH,Nb,n,nn,k,xgrid,
starchdpmin,...
    starchmindp),tgrid,[p_init;pby_init;pba_init;enz_init;
Ey0_init;...
    Ea0_init;EtOH_init;c_init;],options);

```

```

function dydt=combined_pbmcyber_ode_ulgen(t,y,kla,kminusla,...
    k2a,kly,kminusly,k2y,MEy,MEa,MEtOH,Nb,n,nn,k,xgrid,...
    starchdpmin,starchmindp)

K(1)=k(1);
K(2)=3*K(1);
K(3)=3*K(2);
K(4)=k(2);
K(5)=k(3);
Ks(1)=k(4);
Ks(2)=3*Ks(1);
Ks(3)=3*Ks(2);
Ks(4)=k(5);
Ks(5)=k(6);
yld(1)=k(7);
yld(2)=0.95*yld(1);
yld(3)=0.95*yld(2);
mu(1)=k(8);
mu(2)=0.88*mu(1);
mu(3)=0.88*mu(2);
mu(4)=k(9);
mu(5)=k(10);
ke(1)=k(17);
ke(2)=k(18);
ke(3)=k(19);
ke(4)=k(20);
ke(5)=k(21);
EtOHmaxg=k(22);
yldE(1)=k(23);
yldE(2)=1.05*yldE(1);
yldE(3)=1.05*yldE(2);
cgmax1=k(26);
cgmax2=k(27);
beta_E1=k(28);
beta(1)=k(29);
beta(2)=beta(1);
beta(3)=beta(1);
beta(4)=k(30);
beta(5)=k(31);
Kd=k(32);

%state variables
g=zeros(Nb,1); %substrates (mol/L)
gby=zeros(Nb,1); %enz complexes by glucoamylase (mol/L)
gba=zeros(Nb,1); %enz complexes by alpha amylase (mol/L)
for i=1:Nb
    g(i,1)=y(i); %soluble substrate - DP(i)
    gby(i,1)=y(Nb+i); %bounded substrate by glucoamylase - DP(i)
    gba(i,1)=y(2*Nb+i); %bounded substrate by alpha-amylase - DP(i)
end
sugar(1,1)=(g(1)+gby(1)+gba(1))*(162*xgrid(1)+18); %total DP1 (g/L)

```

```

sugar(2,1)=(g(2)+gby(2)+gba(2))*(162*xgrid(2)+18); %total DP2 (g/L)
sugar(3,1)=(g(3)+gby(3)+gba(3))*(162*xgrid(3)+18); %total DP3 (g/L)

starch=0; %total starch (g/L)
for i=starchdpm:Nb
    starch=starch+(162*xgrid(i)+18)*(g(i)+gby(i)+gba(i));
end

reducing_sugar=0; %total reducing sugars (g/L)
for i=starchmindp:starchdpm-1
    reducing_sugar=reducing_sugar+(162*xgrid(i)+18)*(g(i)+gby(i)+gba(i));
end

enz(1,1)=y(3*Nb+1); %dimensionless enzyme level - DP1
enz(2,1)=y(3*Nb+2); %dimensionless enzyme level - DP2
enz(3,1)=y(3*Nb+3); %dimensionless enzyme level - DP3
Ey0=y(3*Nb+4); %total extracellular glucoamylase (g/L)
Ea0=y(3*Nb+5); %total extracellular alpha-amylase (g/L)

EtOH=y(3*Nb+6); %ethanol (mol/L)
Ethanol_0=EtOH*MEtOH; %ethanol (g/L)

c=y(end); %biomass conc (g/L)

re=zeros(5,1); %enz synthesis rate (h^-1)
r=zeros(5,1); %biomass growth rate (g L^-1 h^-1)
for i=1:3
    re(i,1)=ke(i)*sugar(i)/(Ks(i)+sugar(i));
    r(i,1)=mu(i)*enz(i)*sugar(i)*c/((K(i)+sugar(i))*...
        *(1+(Ethanol_0/EtOHmaxg)));
end

re(4,1)=ke(4)*(starch+reducing_sugar)*c/(...
    Ks(4)+(starch+reducing_sugar)); %extracellular glucoamylase ✓
induction
re(5,1)=ke(5)*(starch+reducing_sugar)*c/(...
    Ks(5)+(starch+reducing_sugar)); %extracellular alpha-amylase ✓
induction

r(4,1)=mu(4)*Ey0*(starch+reducing_sugar)/...
    ((K(4)+(starch+reducing_sugar))*(1+Ethanol_0/EtOHmaxg))*...
    (1+sugar(1)/cgmax1)); %growth indicator for glucoamylase ✓
production
r(5,1)=mu(5)*Ea0*(starch+reducing_sugar)/...
    ((K(5)+(starch+reducing_sugar))*(1+Ethanol_0/EtOHmaxg))*...
    (1+sugar(1)/cgmax2)); %growth indicator for alpha-amylase ✓
production

%cybernetic variables
u=r./sum(r); %dimensionless

```

```

v=r(1:3)./max(r(1:3)); %dimensionless

%enzyme balance
Ey=Ey0/MEy-sum(gby);
Ea=Ea0/MEa-sum(gba);

%ODEs
dg=zeros(1,Nb);
dgby=zeros(1,Nb);
dgba=zeros(1,Nb);
denz=zeros(1,3);

%biomass (g L-1 h-1)
dc=sum(r(1:3).*v(1:3))-Kd*c;

%ethanol (mol L-1 h-1)
dEtOH=sum((yldE(1:3)'./yld(1:3)')*.r(1:3).*v(1:3))/MEtOH;

%DP1
dg(1)=-kly(1)*Ey*g(1)+kminusly(1)*gby(1)+k2y(2)*gby(2)...
+sum(k2y(2:end).*gby(2:end))-kla(1)*Ea*g(1)+kminusla(1)...
*gba(1)+sum(nn(1,2:end)'.*k2a(2:end).*gba(2:end))...
-(r(1)*v(1)/yld(1))/(180); %soluble DP1 (mol L-1 h-1)
dgby(1)=kly(1)*Ey*g(1)...
-kminusly(1)*gby(1); %bounded DP1 by gluco (mol L-1 h-1)
dgba(1)=kla(1)*Ea*g(1)...
-kminusla(1)*gba(1); %bounded DP1 by alpha (mol L-1 h-1)
denz(1)=re(1)*u(1)-beta(1)*enz(1)-(enz(1)/c)*dc; %enzyme level (h-1)

%DP2 - DP3
for i=2:3
    summation=0;
    for j=i:Nb
        summation=summation+n(i,j)*k2y(j)*gby(j);
    end
    summation1=0;
    for jj=i+1:Nb
        summation1=summation1+nn(i,jj)*k2a(jj)*gba(jj);
    end
    dg(i)=-kly(i)*Ey*g(i)+kminusly(i)*gby(i)...
-kla(i)*Ea*g(i)+kminusla(i)*gba(i)+summation+summation1...
-(r(i)*v(i)/yld(i))/(162*xgrid(i)+18);
    dgby(i)=kly(i)*Ey*g(i)-kminusly(i)*gby(i)-k2y(i)*gby(i);
    dgba(i)=kla(i)*Ea*g(i)-kminusla(i)*gba(i)-k2a(i)*gba(i);
    denz(i)=re(i)*u(i)-beta(i)*enz(i)-(enz(i)/c)*dc;
end

%change in extracellular glucoamylase concentration (g L-1 h-1)
dEy0=re(4)*u(4)-beta(4)*(1+beta_E1*Ethanol_0^2)*Ey0;

%change in extracellular alpha amylase concentration (g L-1 h-1)

```

```

dEa0=re(5)*u(5)-beta(5)*Ea0;

%DP4 - DPN
for i=4:Nb
    summation=0;
    for j=i:Nb
        summation=summation+n(i,j)*k2y(j)*gby(j);
    end
    summation1=0;
    for jj=i:Nb
        summation1=summation1+nn(i,jj)*k2a(jj)*gba(jj);
    end
    dg(i)=-kly(i)*Ey*g(i)+kminusly(i)*gby(i)...
        -kla(i)*Ea*g(i)+kminusla(i)*gba(i)+summation+summation1;
    dgby(i)=kly(i)*Ey*g(i)-kminusly(i)*gby(i)-k2y(i)*gby(i);
    dgba(i)=kla(i)*Ea*g(i)-kminusla(i)*gba(i)-k2a(i)*gba(i);
end

%output the state variables
dydt=[dg';dgby';dgba';denz';dEy0;dEa0;dEtOH;dc;];

```